12-2016

Exploring the Influence of Urban Watershed Characteristics and Antecedent Climate on In-Stream Pollutant Dynamics

Laurel Elizabeth Christian
University of Tennessee, Knoxville, lchrist4@vols.utk.edu

To the Graduate Council:

I am submitting herewith a thesis written by Laurel Elizabeth Christian entitled "Exploring the Influence of Urban Watershed Characteristics and Antecedent Climate on In-Stream Pollutant Dynamics." I have examined the final electronic copy of this thesis for form and content and recommend that it be accepted in partial fulfillment of the requirements for the degree of Master of Science, with a major in Environmental Engineering.

Jon Hathaway, Major Professor

We have read this thesis and recommend its acceptance:

John Schwartz, Ana Szynkiewicz

Accepted for the Council:
Carolyn R. Hodges

Vice Provost and Dean of the Graduate School

(Original signatures are on file with official student records.)

Exploring the Influence of Urban Watershed Characteristics and Antecedent Climate on In-Stream Pollutant Dynamics

A Thesis Presented for the
Master of Science
Degree
The University of Tennessee, Knoxville

Laurel Elizabeth Christian
December 2016

ACKNOWLEDGEMENTS

First, I would like to thank my adviser, Dr. Jon Hathaway, for his support and encouragement over the last two and a half years. His guidance helped me throughout this project and writing of this thesis. In addition, I would like to thank my research group. Whether it was taking a coffee break, getting a second opinion on projects, or helping each other out in the field, I'm thankful for their time. I would also like to thank my family whose unwavering support throughout my years of studies means so much to me. I could not have accomplished all that I have without them.

ABSTRACT

Understanding pollutant fate and transport in urban watersheds is a challenging endeavor, as heterogeneity of land use, precipitation patterns, and pollutant loadings add complexity to the system. As a result, many currently utilized water quality models exhibit poor performance. One main challenge in urban system modeling is the lack of quality data sets for model development, calibration, and testing, resulting in the need for high quality data collection. Although recent studies have begun to investigate pollutant transport in urban watersheds to aid these models, these studies have focused primarily on the end-of-pipe as the point of interest (i.e. prior to stormwater entering an open stream channel). However, it is likely that in-stream processes will influence pollutants leaving urban watersheds when the system is viewed at a larger scale. The results of the high-resolution sample collection performed in this study improve our understanding of water quality trends at the watershed scale (including the influence of in-stream processes), and ultimately can be used to improve urban watershed water quality models. Studies have shown correlations between runoff water quality variability and land use classification, antecedent climate, and rainfall factors. Further investigation in urban streams could support these parallel relationships. The goals of the study include: (1) identify trends in water quality due to watershed characteristics using two inter-storm analysis methods, and (2) report the effects of explanatory factors on these trends in Knoxville, TN, streams. Ultimately, this study intends to contribute to water quality prediction during storm events by analyzing how watershed characteristics, rainfall patterns, flow regimes, and antecedent climate factors affect water quality.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

1 INTRODUCTION

Urban stormwater runoff is one of the major contributors to stream water quality degradation. Multiple pollutants can be found in urban runoff, such as sediment, bacteria, oil and grease, metals, nutrients, and other harmful toxins ("Nonpoint source: Urban Areas", 2016). Since the implementation of the U.S. Environmental Protection Agency (EPA) in 1970, and the Clean Water Act in 1972, many researchers have focused on urban stormwater effects on water bodies as nonpoint sources of pollution. These studies commonly evaluate pollutant transport during storm events based on the proclivity for the first flush. The first flush is observed when the majority of the pollutant load is delivered in the beginning of a storm event (Gupta & Saul, 1996). Recognizing and understanding the first flush is important when designing stormwater control measures (SCMs). SCMs are designed to provide control of flow and pollution by capturing the first few inches of runoff volume (Deletic, 1998). Despite this critical need, definitions as to how much pollutant load is delivered, and therefore treated, in the first amount of runoff volume are unclearly defined in literature. Further, the propensity a pollutant has for the first flush can be used to compare differences in transport amongst pollutants (Hathaway et al. 2012). Despite the large number of first flush studies published, one best method has not been agreed upon. For instance, a new method by Bach, McCarthy, and Deletic (2010) has solved some uncertainties in typical first flush definitions; however, only a few studies have used this new method. Regardless, researchers are still redefining and deciding on the best definition for the first flush phenomenon to allow its broad application over multiple contaminants and utilizing these data to determine the parameters affecting it. The various methodologies employed to date are described below in detail.

The most common analysis used to determine the first flush effect is the normalized cumulative curve. This curve is defined by plotting the normalized cumulative pollutant mass (M') at time, t, versus the normalized cumulative runoff volume (V') at time, t. M' and V' are calculated by dividing the cumulative mass, m(t), and cumulative volume, v(t), by total pollutant load and total volume, respectively, for the entire runoff event as shown:

$$M' = \frac{m(t)}{M}$$

$$V' = \frac{v(t)}{V}$$

where M = total pollutant mass over storm event and V = total volume of flow over storm event.

The curve shows the fractional amount of pollutant in relation to the fractional amount of volume draining from a watershed or catchment during an event. Initially, researchers defined the first flush as occurring when the M'V' curve was above the 45-degree bisector during the event (that is, when m(t)>v(t)). When the curve lies below the bisector, it suggests a dilution effect. Since this initial definition, many researches have restricted this definition in different ways. For example, Geiger (1984) suggested the first flush is the volume before the largest gap between M'V' curve and the bisector, and that this gap must be greater than 0.2. In other studies, a power function is fit to the curve where the exponential coefficient, b, represents the distance from the curve to the 45-degree bisector (Saget, Chebbo, & Bertrand-Krajewski, 1996).

$$M' = V'^{b}$$

With this coefficient, various explanatory variables can be investigated for their influence on the first flush. The power function uses b to represent the strength of the first flush. The lower the value of b, the stronger the first flush of an event. Using the b coefficient, Batroney, Wadzuk, and Traver (2010) determined a first flush existed for total suspended solids (TSS), nitrate, chloride, copper, and cadmium for an impervious parking lot on Villanova University's campus in Philadelphia, Pennsylvania. However, based on the multiple studies that use the M'V' curve in the variations described above (Bertrand-Krajewski, Chebbo, & Saget, 1998; Gupta & Saul, 1996; Lee & Bang, 2000; Lee, Bang, Ketchum, Choe, & Yu, 2002; Lee, Yu, Bang, & Choe, 2003; Saget et al., 1996; Sansalone & Buchberger, 1997; Sansalone & Cristina, 2004; Sansalone, Koran, Smithson, & Buchberger, 1998; Taebi & Droste, 2004), substantially variable results may be found between storms and locations. Because of this, Bertrand-Krajewski et al. (1998) suggest further statistical and site-specific analysis is necessary to further understand the first flush phenomenon.

Another common method many researchers have used to define the first flush is by applying a threshold, that is, defining a specific percentage of the total event volume that has passed at which the first flush should be evaluated. In some cases, this results in a binary decision of whether or not a first flush has occurred by somewhat arbitrarily defining a percent load that must have passed at the threshold volume. Some common thresholds are 50% of load is delivered in the first 25% runoff volume (Flint & Davis, 2007), 80% of load is delivered in the first 20% runoff volume (Sansalone & Cristina, 2004), and 80% of load is delivered in the first 30% of runoff volume (Bertrand-Krajewski et al., 1998). In other studies, the strength of the first flush was defined as the percentage of total load delivered in the first 30% of event volume, or FF_{30}. Many recent studies in literature have used this definition (Bach et al., 2010; Hathaway & Hunt, 2011; Hathaway, Winston, Brown, Hunt, & McCarthy, 2016; Li, Yin, He, & Kong, 2007).

Finally, a fundamentally different approach was proposed by Bach et al. (2010). The authors point out that conventional first flush definitions do not take into account the size of the storm event, with each event being analyzed individually and given equal weight instead of being compiled and viewed collectively. Thus, the first flush is equally investigated in small and large events despite the possibility that larger runoff volumes may more readily deplete surface sources throughout the course of a storm. Conversely, the methodology by Bach et al. (2010) allows for differences in pollutant transport as runoff depth increases. The new proposed methodology separates runoff from each event into "slices" of a selected depth. As these slices represent the same runoff depth across all storms, they can be compiled and used to compare trends across multiple events. This new method gives a more definitive definition of the first flush phenomenon by treating the first flush as a site characteristic rather than on that changes event-by-event. Hathaway et al. (2016) utilized this methodology when analyzing temperature patterns in urban runoff from multiple catchments, concluding that for some pollutants, in this case temperature, this methodology appears more appropriate.

Common contaminants evaluated for the first flush phenomenon in urban stormwater runoff are sediments, indicator bacteria, oil and grease, metals, nutrients, pH, and temperature (Bertrand-Krajewski et al., 1998; Deletic, 1998; Hathaway & Hunt, 2011; Hathaway, Tucker, Spooner, & Hunt, 2012; Hathaway et al., 2016; Lee et al., 2002; Sansalone & Buchberger, 1997; Sansalone & Cristina, 2004). Sediment is the most common constituent to exhibit a first flush. Deletic (1998); Lee and Bang (2000); Lee et al. (2002); Lee et al. (2003); Sansalone and Buchberger (1997); Taebi and Droste (2004); and Hathaway and Hunt (2011) have all observed flushing of sediments over multiple events, albeit mostly weak and inconsistent. It is important to note most of the first flush results are using various first flush definitions; however, observations of the first flush for most contaminants are weak. Hathaway et al. (2016) is the only study to have repeated the Bach et al. (2010) methodology and provided detection of a thermal first flush throughout multiple events from one catchment. Future studies using this methodology might result in more insight into the first flush phenomenon; in particular, this method may lead to observation of stronger first flush effects for pollutants.

Additional factors, such as watershed characteristics, antecedent dry weather period, sewer system conditions, amount of accumulated pollutants over a catchment, rainfall intensity, and storm size have been suggested in literature as factors affecting the first flush. However, the complex nature of the first flush makes the identification of impactful explanatory variables difficult. Furthermore, past research studies have collected samples from catchment outfalls prior to runoff entering an open channel stream system. If stormwater samples were collected from in-stream, where new processes dominate transport, pollutant load patterns will likely be different. The few studies that have collected samples from in-stream did so manually, at timed intervals, and typically did not test for numerous water quality parameters (Characklis & Wiesner, 1997). Lee et al. (2002) took samples in-stream and from manholes, and compared both without considering in-stream processes. Overall, the study's results appeared to be highly variable for pollutants in both residential and industrial areas. Studies have found the top layer of the streambed is re-suspended during storm flow, potentially having a large impact on the in-stream first flush (Pachepsky, Guber, Shelton, & Hill, 2009; Surbeck, Shields, & Cooper, 2016). Re-suspension is most likely due to event intensity and runoff depth. Storm flow increases water velocities which results in re-suspension of sediments from the stream bank and bed. As shown in literature, bacteria and other pollutants are sometimes transported while attached to sediments (Characklis et al., 2005; Krometis et al., 2007). Additionally, the streambed has been found to be a major source for *Escherichia coli* during events without input from stormwater runoff. Muirhead, Davies-Colley, Donnison, and Nagels (2004) performed an artificial flood event controlled by a dam valve and found high concentrations of *E. coli* coming from the stream channel. Surbeck et al. (2016) also observed a first flush due to the re-suspension of the top layer of their simulated stream sand bed experiment. The study's procedure, which is intended to mimic a natural stream, determined substantial concentrations of fecal indicator bacteria is found in streambed sediments and the top layer of the streambed contributes to the first flush. Slow infiltration from sewer leaks has also been suggested to cause this pattern seen in-stream (Bach et al., 2010; D. McCarthy, 2009).

Consequently, it is important to consider in-stream pollution sources contributing to a stream's health, not only point sources from stormwater outfalls.

Ultimately, this understanding should enhance water quality models by informing our understanding of pollutant fate and transport in urban watersheds. A review of surface water quality models concluded models need to be standardized through more research and investigation of water quality conditions (Wang, Li, Jia, Qi, & Ding, 2013). Model verification and calibration is an essential step when designing a water quality model. One study found a few current water quality models performed poorly when tested against monitored data. To improve model performance, they suggested further investigation of the relationship between pollutants and their explanatory factors (Dotto et al., 2010). Sample collection from stormwater outfalls does not give insight into pollutant dynamics within the stream, which is often the point of interest for water quality models.

Understanding pollutant fate and transport in urban watersheds is an area of ongoing need that can be greatly informed by studies which characterize patterns of pollutant export. As aforementioned, the majority of previous studies collected samples from outfalls before stormwater reached an open channel stream system. As recent literature suggests, in-stream processes, i.e. bedform, re-suspension (of both sediment and pollutants held therein), and bank erosion, could strongly influence the first flush of multiple constituents. This study intends to investigate various constituents' transport patterns during a storm event from three different in-stream locations in Knoxville, Tennessee. Using a traditional first flush definition and the new slice method introduced by Bach et al. (2010), each pollutant will be evaluated for the first flush phenomenon. Further, the inter-event variation in first flush strength will be compared to explanatory variables such as watershed attributes and antecedent climate.

2 MATERIALS AND METHODS

2.1 Study Area

The project locations for this study are Second Creek, Third Creek, and Williams Creek, three predominately urbanized streams in downtown Knoxville, TN (Figure 1). Each monitoring station collects samples from a reach of natural open channel within a given stream. The watersheds vary in size from 564.3 hectares to 4212.9 hectares, and impervious surface percentages range from 26% to 45%.

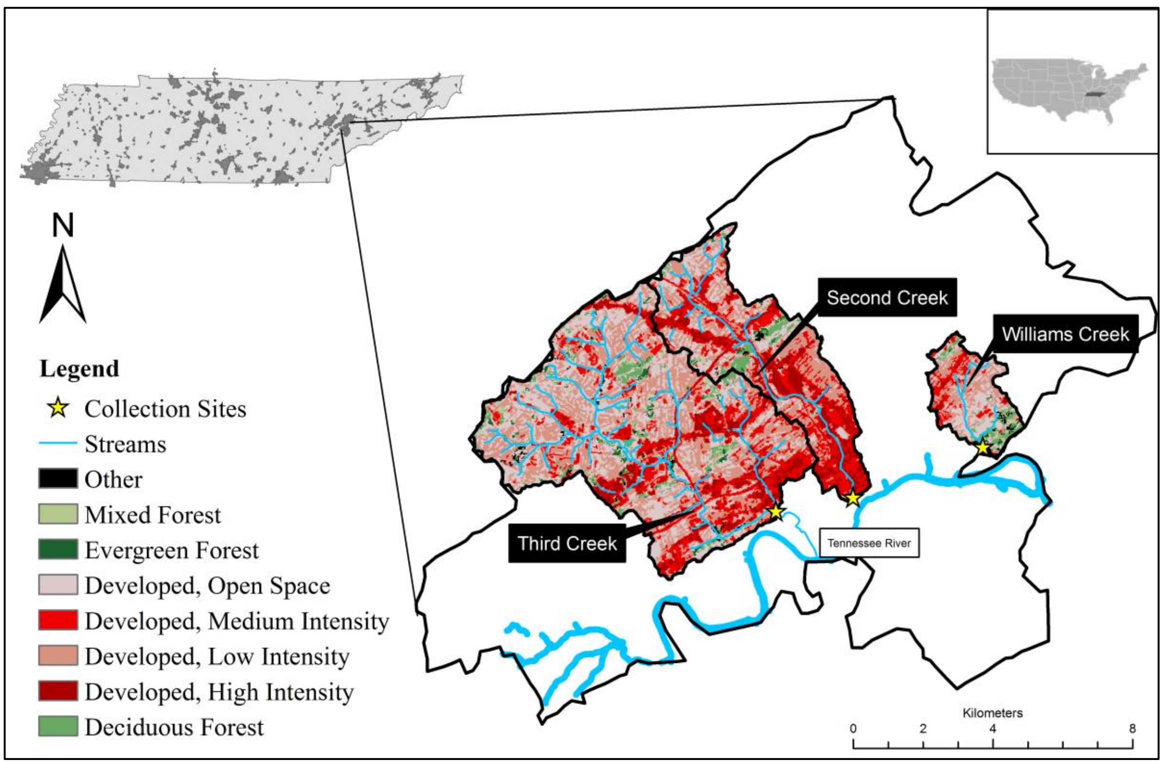


Figure 1. Study areas in Knoxville, TN with the 2011 National Land Cover Data

Each watershed has similar land use attributes, primarily open space and low to medium density development, with separate waste water and stormwater systems (Table 1). Second Creek has received attention in the past through a study investigating its geomorphic response to urbanization, with results indicating that at the time of the study, Second Creek was still adjusting to the increase of stormwater runoff due to impervious surface increase (Grable & Harden, 2006).

Table 1. Area, percent impervious, land use, number of events, and runoff coefficient for three site locations.

Site	Watershed Area (ha)[a]	Percent Impervious[b]	Land Use	No. of Events	Runoff Coefficient (ROC)[c]
Second Creek	1800.27	45	Developed, Low Intensity – 33% Developed, Medium Intensity – 29% Developed, Open Space – 18%	17	0.24
Third Creek	4212.90	33	Developed, Low Intensity – 37% Developed, Open Space – 31% Developed, Medium Intensity – 17%	7	0.08
Williams Creek	564.27	26	Developed, Low Intensity – 35% Developed, Open Space – 32% Developed, Medium Intensity – 15%	8	0.63

[a] values calculated from 30 m DEM raster obtained from U.S. Geological Survey (USGS)
[b] values calculated from impervious cover from the National Land Cover Database 2011 (NLCD)
[c] values calculated based on collected storm volume and rainfall depth data and dependent on each watershed area

Second Creek's substrate is composed of sand, silt, and clay regolith with a dominate presence of gravel and cobble (Grable & Harden, 2006). Similar observations were made regarding Third Creek and Williams Creek substrate. The soils throughout all three watersheds are silt loam or silty clay loam. However, William's Creek also contains a large amount of pits, mines, and dumps within the catchment (SSURGO, 2014). Each stream was on the United States Environmental Protection Agency's 303d stream impairment list in 2012. Second Creek, Third Creek, and Williams Creek are impaired by collection system failures and discharges from municipal separate storm sewer systems (MS4) areas. This causes Second Creek and Third Creek to have high levels of nitrate/nitrite and suffer loss of biological integrity due to siltation. All three creeks have habitat alterations due to anthropogenic activities, and contain high levels of *Escherichia coli* (303d list).

2.2 Sampling Methodology

The Second Creek monitoring station was installed in the summer of 2014; Third Creek and Williams Creek stations were installed in the spring of 2015. The monitoring equipment for Second Creek consists of an ISCO 310 (i.e. an ultrasonic level recorder) which is mounted under a pedestrian bridge and feeds stream level readings to an ISCO Signature Flow Meter. The Signature uses a stage-discharge relationship developed for the site (see Appendix A) to convert these readings to flow, and sends flow paced pulses to an Avalanche refrigerated auto-sampler for sample collection. The monitoring equipment for the other sites, Williams Creek and Third Creek, consists of an ISCO 4230 which records level, converts these readings to flow using site-specific stage-discharge relationships, and sends flow-paced pulses to an ISCO 3700 auto-sampler.

The stage discharge relationship for Second Creek was developed by matching cross section data gathered by the University of Tennessee with historical readings from a level-velocity meter. Third Creek and Williams Creek stage-discharge relationships were determined based on historical flow data recorded by the City of Knoxville. The level offsets between the City of Knoxville's in-stream installation and this study's in-stream installation was applied to the stage-discharge relationship for each location to ensure accurate flow readings. The stage-discharge relationship for each location can be found in the Appendix.

Flow data was recorded every 5 minutes at each site to well-characterize the hydrographs of the flashy streams. Samplers were triggered by a rise in water level during targeted storms events to avoid collection of baseflow. Each sample bottle contained four aliquots and was analyzed discretely. Sample bottles were sterilized by submerging them in a hydrochloric acid bath for 30 minutes. Next, they were rinsed with deionized water, and autoclaved at 121 °C for 20 minutes. The acid bath was used to minimize cross-contamination of trace metals and the autoclave was used to minimize cross-contamination of microbes. In addition, two tipping bucket rain gauges were installed and used for the Second Creek site. Rainfall data were collected at 1 minute increments. On-site manual rain gages were also installed to allow validation of tipping bucket data. Other rainfall data used for Williams Creek and Third Creek were obtained through City of Knoxville's Stormwater Engineering Division.

2.3 Sample Analysis

Lab analyses and/or sample preservation techniques were performed within 24 hours of sample collection. Samples were analyzed for fecal coliform, *E. coli*, TSS, Cu^{2+}, and NO_3^-. The Colilert (IDEXX Laboratories, Westbrook, Maine) analysis was adjusted per Yakub et al. (2002) to allow enumeration of fecal coliform and *E. coli*. Samples underwent either a 100:1 or a 1000:1 dilution (depending on bacteria counts and season) to avoid errors stemming from concentrations exceeding the maximum thresholds of the analysis. Samples were incubated for 4 hours at 37ºC then switched to a 44.5ºC incubator for 20 hours (Yakub et al., 2002). There were only four occurrences, all from Williams Creek, when bacteria counts were above or below testing detection limits (2 above, 2 below). In such cases, the minimum or maximum detection limits, respectively, were used in the analyses similar to the method utilized by McCarthy et al. (2012). Total suspended solids (TSS) was measured using the SM 2540 D filtration method (APHA, 2005). Last, samples were filtered and placed in refrigeration for the IC (Ion Chromatography, Method 300.1 – anions and cations) and ICP-AES (Inductively Coupled Plasma-Atomic Emission Spectrometry, Method 200.7 – trace metals) analyses to determine Cu^{2+} and NO_3^-. IC analyses were run within the 28-day hold time. The ICP designated sample was preserved by a dose of nitric acid and the analyses were run within the 6-month hold time.

2.4 Data Analysis

Since samples are collected in-stream, start and end times of each event had to be delimited to differentiate from baseflow. The start of each storm was defined as the time just before the beginning of the rising limb of the hydrograph, that is, when flows increased above baseflow. The end of the event was identified as the intersection of the tangents from the recession limb and the receding limb of the hydrograph. Using this procedure for each storm event allowed for consistency in storm flow duration determination. Baseflow was not separated from stormwater runoff because it is assumed that mixing occurs within the stream; therefore, the samples collected represent these mixed concentrations.

2.4.i FF_{30}: Traditional First Flush Analysis

The normalized cumulative runoff volume (V') versus cumulative pollutant mass (M') was calculated for each storm event and location. The curve shows the fractional amount of pollutant in relation to the corresponding fractional amount of volume draining from a watershed area during a given event. A flushing effect is observed when the M'V' curve is above the 45-degree bisector during the event or when $m(t)>v(t)$. When the curve lies below the 45-degree bisector, the plot suggests a dilution effect. The strength of the first flush is defined herein as the percentage of total pollutant load delivered in the first 30% of event volume, or FF_{30}. While a variety of different thresholds have been used, this is consistent with many recent studies in literature, allowing comparison of results (Bach et al., 2010; Hathaway & Hunt, 2011; Hathaway et al., 2016; Li et al., 2007).

2.4.ii Bach Slice Method: Modified First Flush Analysis

A new definition used to determine the first flush was introduced by Bach et al. (2010). To perform this assessment, each storm event's cumulative runoff volume was first converted into cumulative runoff depth. This was achieved by dividing the event volume by the product of the runoff coefficient and watershed area (Table 1). The runoff coefficient was calculated for each watershed by dividing each storm event depth by its rainfall depth and averaging the resultant fractions. Next, each storm event is broken up into slices of runoff depth. For each slice, the pollutant concentrations are interpolated for each event to represent the slice's average pollutant concentration. Using the Wilcoxon Rank Sum test and a 5% level of statistical significance, slices are grouped together based on statistical similarities. The runoff depth was decided on the data availability from the catchment's number of storms collected, and the slice analysis was only performed when there were at least five storm event samples within the slice. Further, the pollutants concentrations were normalized by dividing by the event mean concentration of each event. This was to ensure each event's values were presented on a common scale. Adjacent grouping occurs until the next slice of statistical difference is found; then a new group is started. If all the slices from one site location can be grouped into one large slice, then the first flush did not occur. The first flush did occur if there is more than one slice group.

2.4.iii Antecedent Climate and Other Explanatory Variables

Temperature and relative humidity data was used from Knoxville Tyson-McGee Airport obtained using NC CRONOS database (version 2.7.2). All other data was used from site collection data or Knoxville Stormwater Management. Explanatory variables used were chosen based upon previous work done by (Hathaway & Hunt, 2011; D. McCarthy, 2009; D. T. McCarthy, Hathaway, Hunt, & Deletic, 2012). Antecedent condition factors were considered for 2 days and 28 days. The antecedent dry weather period (ADWP) factor was considered for days with less than 0.05, 0.1, and 0.5 in of rainfall.

2.5 Statistical Analysis

Microsoft Excel was used to calculate FF_{30}, incremental slices, and summary statistics. Due to the small sample size and consistently non-normal distributions of the data, non-parametric tests were utilized. The more complex statistical analyses, like

Wilcoxon Ranks Sum Test and Spearman's Rank Correlation, were performed using JMP Pro Version 12. The Wilcoxon Rank Sum test was used to determine statistical difference between slices as instructed by Bach et al. (2010). Additionally, the test was performed to determine statistical significance of FF_{30}. The test reports if the median between FF_{30} values from a watershed is statistically different than 0.3. Other studies also did this or a similar test to determine first flush statistical significance (Bach et al., 2010; Hathaway & Hunt, 2011; D. McCarthy, 2009). The Spearman Rank test was run on each pollutant for Second Creek using various explanatory variables and FF_{30} values. There are not enough events for the sample results to be confident for Third Creek and Williams Creek. Additionally, the Spearman Rank Test was performed to determine if there are any correlations between pollutants. The Spearman Rank Test identifies the strength and direction (r_s) of a relationship between two sets of data (e.g. first flush strength and antecedent climate). P-values are also reported for statistical significance confirmation.

3 RESULTS AND DISCUSSION

Between September 2014 and August 2016, 32 total storm events were collected and tested for pollutants from three different site locations. Each storm was represented by at least five discrete samples, with the maximum number of representative samples being eighteen and the average being nine. As described above, analyses included fecal coliform, *E. coli*, TSS, Cu^{2+}, and NO_3^-. However, two events from Second Creek are missing Cu^{2+} and NO_3^- data, one event from Williams Creek is missing both Cu^{2+} and NO_3^- data, and one event from Third Creek is missing TSS results due to lack of lab resources. Additionally, if 50% of the discrete samples in each storm event were reported as the detection limit, then that event was not used in the analysis for that pollutant. This limitation only occurred for one event at Williams Creek for NO_3^-.

Table 2 shows summary of each creek's water quality data. Overall, median event mean concentrations (EMC) were 12,365 MPN/100ml fecal coliform, 3,583 MPN/100ml *E. coli*, 113 mg/L TSS, 0.0035 mg/L Cu^{2+}, 2.52 mg/L NO_3^- for Second Creek; 11,903 MPN/100ml fecal coliform, 6,123 MPN/100ml *E. coli*, 105 mg/L TSS, 0.0026 mg/L Cu^{2+}, 2.20 mg/L NO_3^- for Third Creek; and 89,033 MPN/100ml fecal coliform, 12,821 MPN/100ml *E. coli*, 124 mg/L TSS, 0.0035 mg/L Cu^{2+}, 2.12 mg/L NO_3^- for Williams Creek. Fecal coliform was found to be the most variable pollutant at each site.

3.1 FF_{30}: Traditional First Flush Analysis

Table 3 shows FF_{30} summary statistics for each contaminant at all three sites. The range of results for FF_{30} vary for pollutants from each site and no pollutant exhibited a

Table 2. Summary EMC statistics: Maximum, minimum, mean, median, and RSD for fecal coliform, *E. coli*, TSS, Cu^{2+}, and NO_3^-

Pollutant	Second Creek					Third Creek					Williams Creek				
	Max	Min	Mean	Median	RSD	Max	Min	Mean	Median	RSD	Max	Min	Mean	Median	RSD
Fecal coliform, MPN/100ml	191,386	2,583	26,476	12,365	170.30	374,920	1,210	69,568	11,903	195.94	757,706	8,680	169,395	89,033	146.89
E. coli, MPN/100ml	13,754	654	5,205	3,726	78.89	22,829	548	8,420	6,123	99.34	54,001	1,933	17,049	12,821	101.40
TSS, mg/L	783	12	183	113	101.97	297	35	132	105	74.09	404	29	179	124	84.48
Cu2+, mg/L	0.0064	0.0022	0.0038	0.0035	34.51	0.0038	0.0015	0.0024	0.0026	34.10	0.0083	0.0017	0.0048	0.0035	64.33
NO3-, mg/L	4.77	0.09	2.46	2.52	53.13	3.82	1.60	2.57	2.20	34.17	3.13	1.56	2.13	2.12	26.40

Table 3. Summary statistics of FF30: Maximum, minimum, mean, median, and RSD for fecal coliform, *E. coli*, TSS, Cu^{2+}, and NO_3^-

Pollutant	Second Creek					Third Creek					Williams Creek				
	Max	Min	Mean	Median	RSD	Max	Min	Mean	Median	RSD	Max	Min	Mean	Median	RSD
Fecal coliform	0.53	0.12	0.30	0.31	40.46	0.87	0.14	0.51	0.46	59.84	0.76	0.16	0.42	0.36	53.15
E. coli	0.60	0.11	0.32	0.32	38.25	0.88	0.25	0.42	0.33	59.63	0.67	0.12	0.40	0.38	53.77
TSS	0.75	0.13	0.47	0.46	37.87	0.55	0.29	0.39	0.39	25.51	0.90	0.24	0.54	0.47	49.40
Cu^{2+}	0.49	0.21	0.36	0.37	19.70	0.46	0.27	0.34	0.32	18.92	0.38	0.24	0.31	0.30	14.81
NO_3^-	0.53	0.15	0.32	0.30	36.10	0.39	0.25	0.30	0.29	16.28	0.55	0.27	0.37	0.32	30.51

first flush for all events. Williams Creek has the greatest TSS variability (as RSD) compared to the other two sites. Catchment size has affected temperature variability (as RSD) (Hathaway et al., 2016), first flush strength (Stein, Tiefenthaler, & Schiff, 2007), and microbe variability (D. McCarthy, 2009) in the past. As an example, Hathaway et al. (2010) found microbes exhibited high variability (as SD) when collected from a small catchment of 5.1 ha. In other literature, first flush strength has typically increased with decreasing catchment size (Lee & Bang, 2000; Lee et al., 2002; Lee et al., 2003; Stein et al., 2007). Characklis and Wiesner (1997) attributed the lack of the first flush from their study to be due to the larger size of their catchment, suggesting larger catchments have a greater affinity to buffer pollutant extremes and variability. A study by Lee and Bang (2000) found strength of first flush to increase with smaller watershed area and greater rainfall intensity. However, Bertrand-Krajewski et al. (1998) found there was not a correlation between first flush strength and small catchment areas. Thus, some variable outcomes have been found in literature in regard to the effect of watershed size on first flush strength. Williams Creek is a smaller watershed than the other two sites, however, in comparison to past study sites, it is still fairly large at 564.2 hectares. Although, TSS FF_{30} mean and median is typically higher at Williams Creek than the other two watersheds, this is not true for all pollutants. So while higher TSS variability (as RSD) is observed at Williams, the strength of multiple pollutants' first flush at the "small" watershed does not support past claims in literature, or at least is not a trend noted when all watersheds are relatively large. It is likely other parameters, for instance percent imperviousness and other land use factors, also affect these trends.

Fecal coliform and *E. coli* are more variable than all other pollutants at each watershed, as seen by their relatively larger RSD, with Third Creek and Williams Creek microbe FF_{30} values being higher than those of Second Creek. Indicator bacteria have been found to have high variability in past research, rarely exhibiting a consistent flushing effect. Attempts have been made to determine which explanatory variables contribute to that variability with antecedent climate suggested as a primary driver (Hathaway, Hunt, & Simmons, 2010; D. McCarthy, 2009). The lowest average RSD was observed for Cu^{2+} and NO_3^-. As expected, TSS FF_{30} values are high for all three watersheds, containing the highest max FF_{30} at Second and Williams Creeks.

Normalized cumulative mass versus volume curves were generated to support findings in Table 3 (Figures 2, 3, and 4). Each line represents a storm event that occurred within a given site. Lines that fall above the 1:1 dotted line suggests some magnitude of first flush occurred. Specifically, for our threshold definition, a first flush occurred if lines are above the 1:1 line at 30% of cumulative event volume. As concluded from Table 3, TSS shows a consistent and stronger first flush than other constituents for all three watersheds with events falling above the 45-degree line on the M'V' curve.

To determine statistical significance of FF_{30}, the Wilcoxon Rank Sum test was performed to determine if the FF_{30} values from each watershed were statistically different than 0.3. The statistical probability from the test, $P_{30\%}$, and the percentage of events with a $FF_{30} > 0.3$ for each pollutant are found in Table 4. Second Creek's *E. coli*, TSS, and Cu^{2+}; Third Creek's TSS and Cu^{2+}; and Williams Creek's TSS median FF_{30} were found to be statistically different than 0.3 with a p-value <0.05 or <0.1 (95% and 90%

Figure 2. Traditional first flush normalized cumulative pollutant load vs cumulative volume curves for Second Creek. (a) fecal coliform, (b) *E. coli*, (c) TSS, (d) Cu^{2+}, and (e) NO_3^-

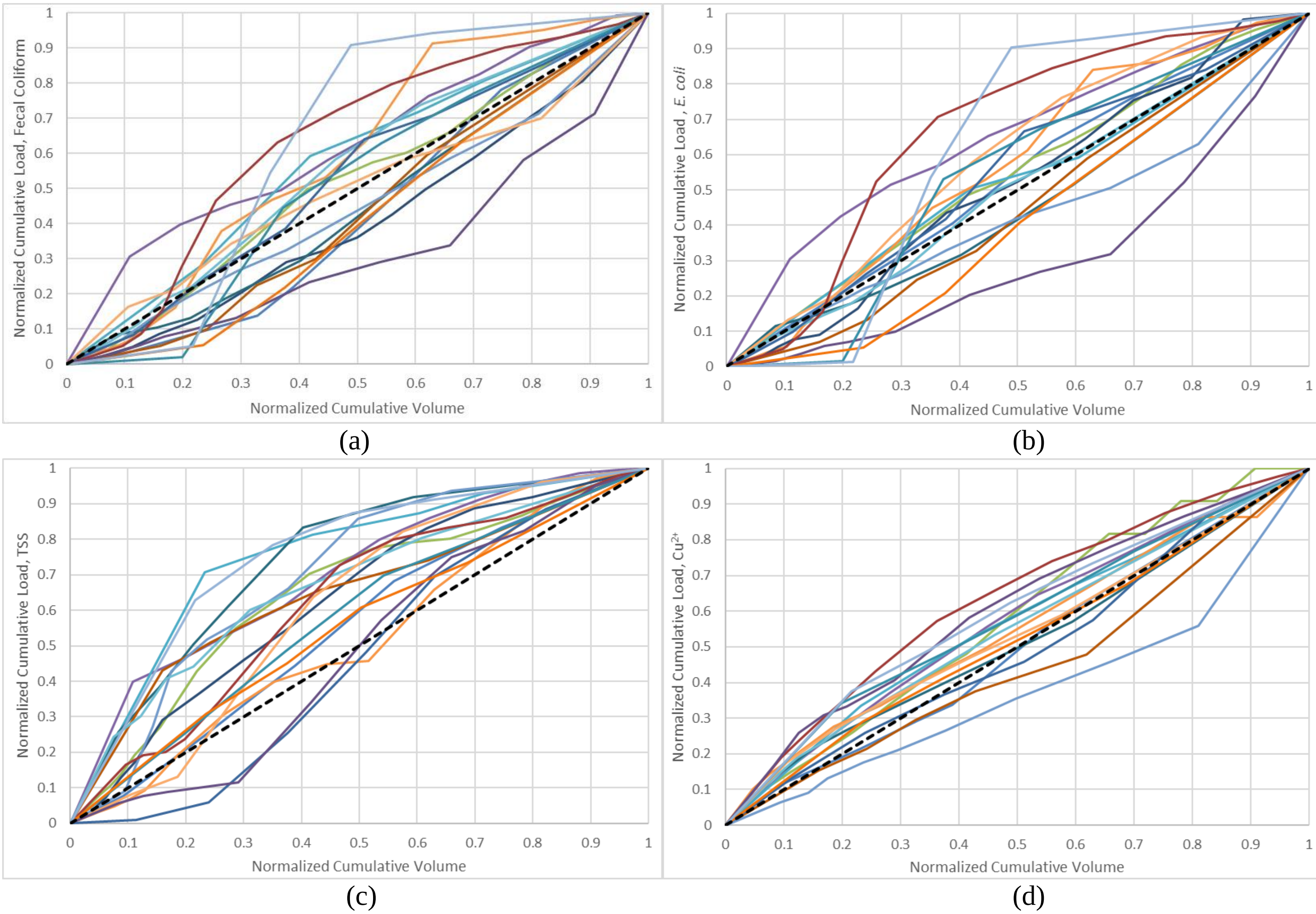
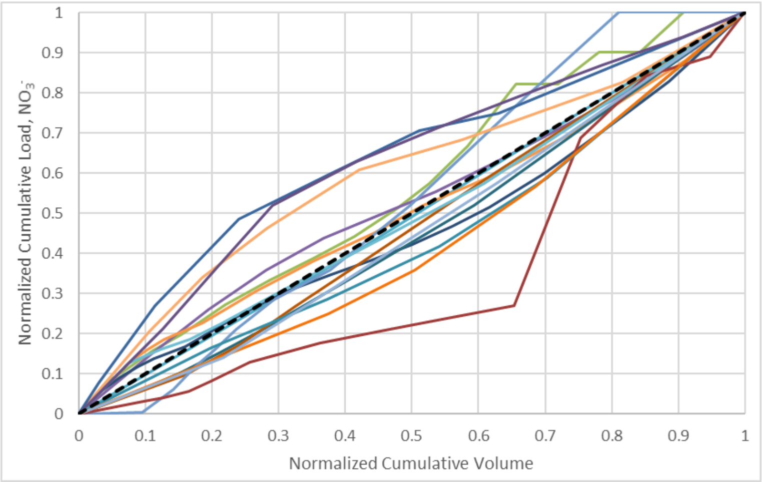


Figure 2 continued



(e)

Figure 2 continued

Figure 3. Traditional first flush normalized cumulative pollutant load vs cumulative volume curves for Third Creek. (a) fecal coliform, (b) *E. coli*, (c) TSS, (d) Cu^{2+}, and (e) NO_3^-

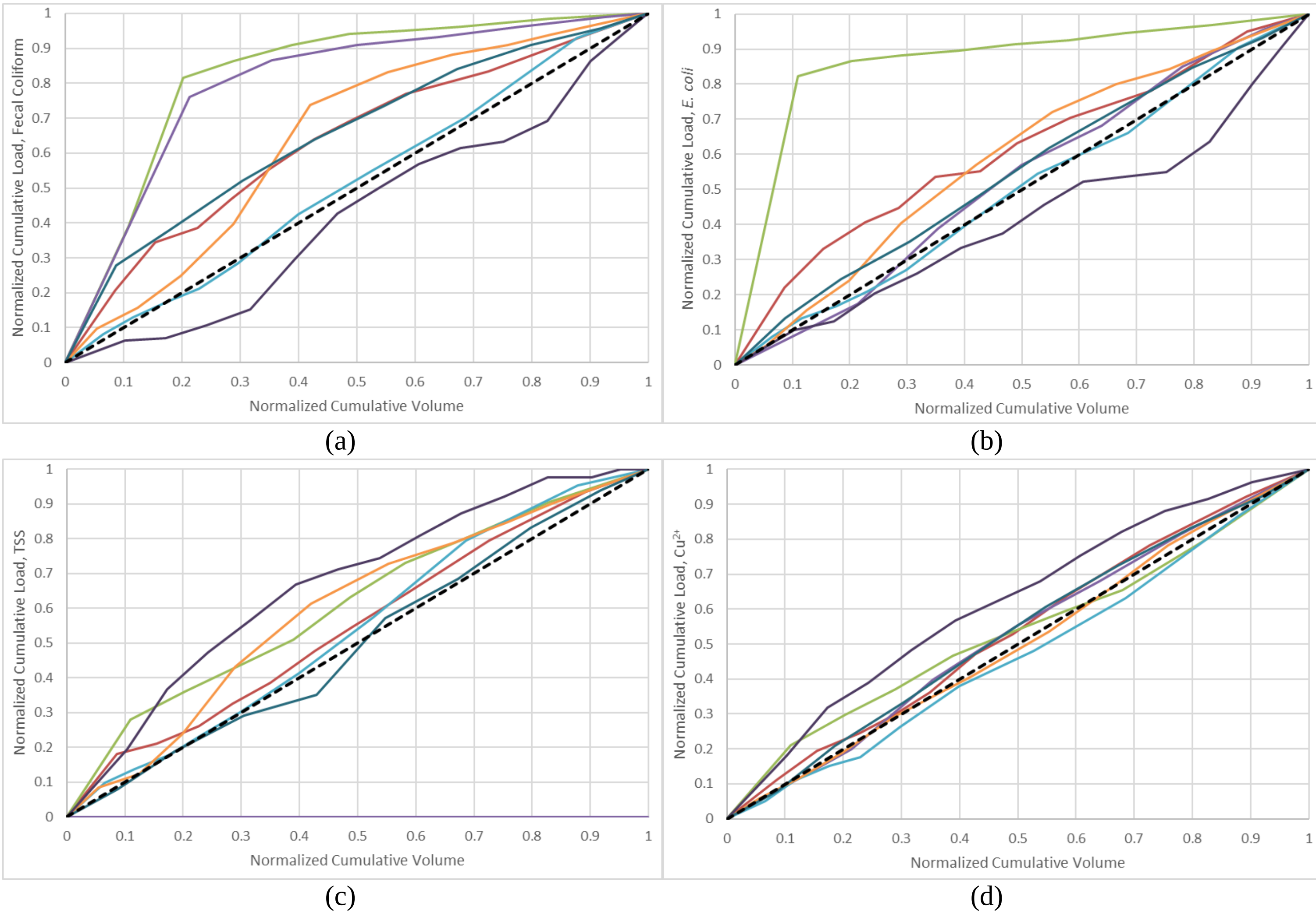
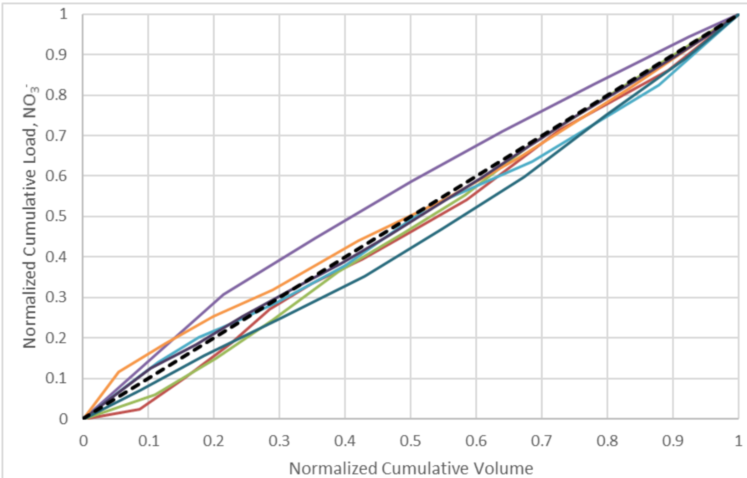


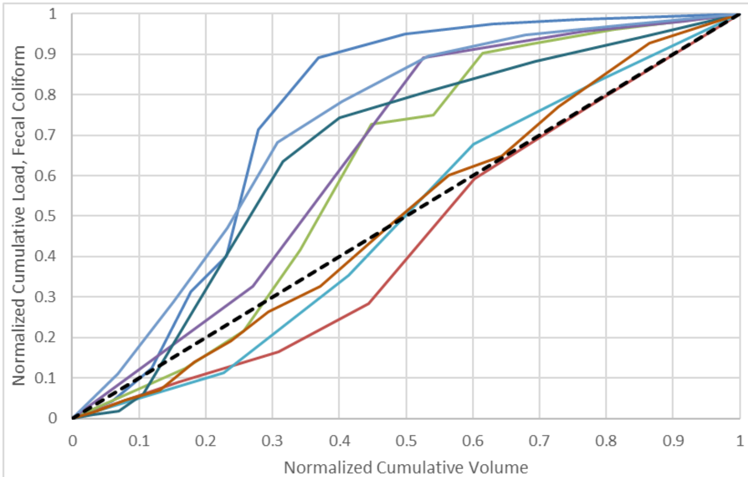
Figure 3 continued



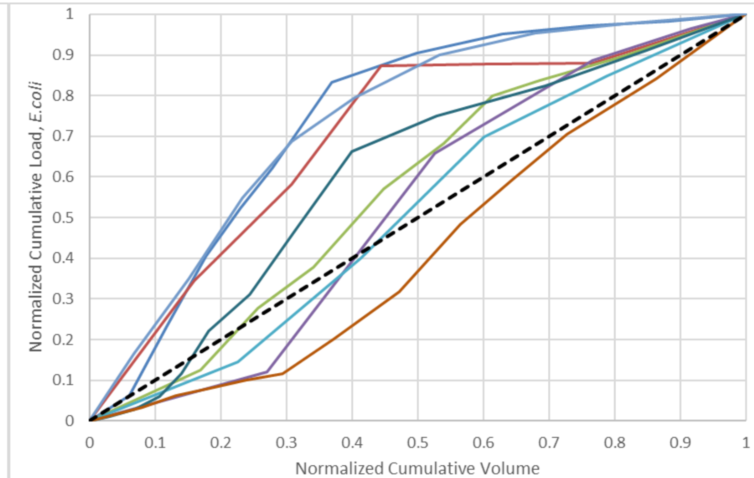
(e)

Figure 3 continued

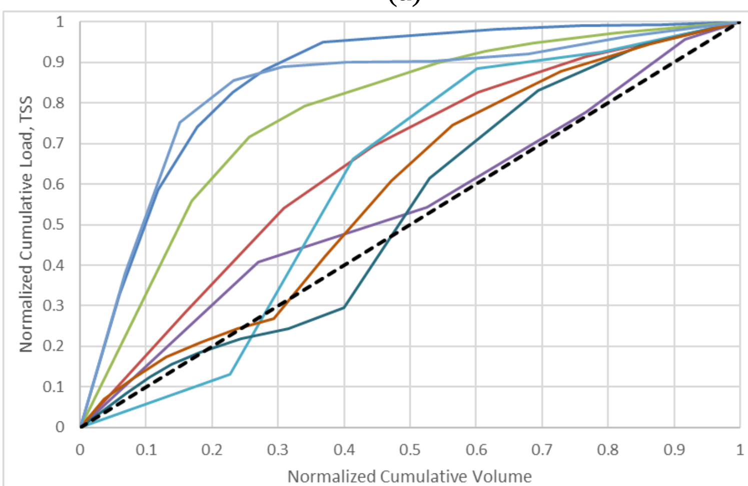
Figure 4. Traditional first flush normalized cumulative pollutant load vs cumulative volume curves for Williams Creek. (a) fecal coliform, (b) *E. coli*, (c) TSS, (d) Cu^{2+}, and (e) NO_3^-



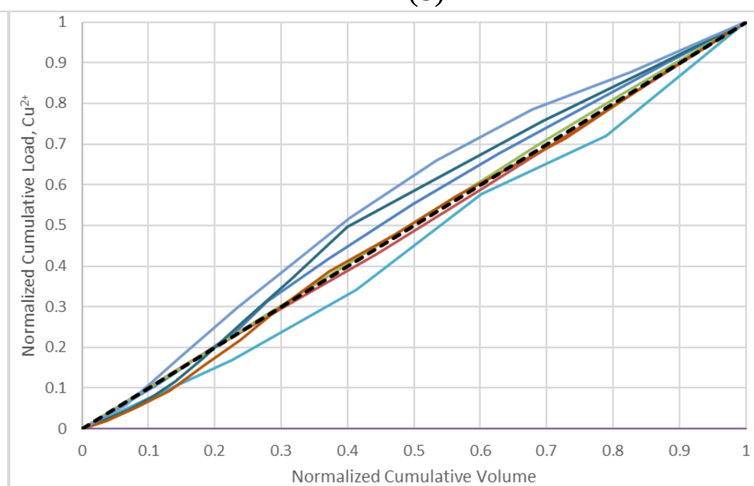
(a)



(b)

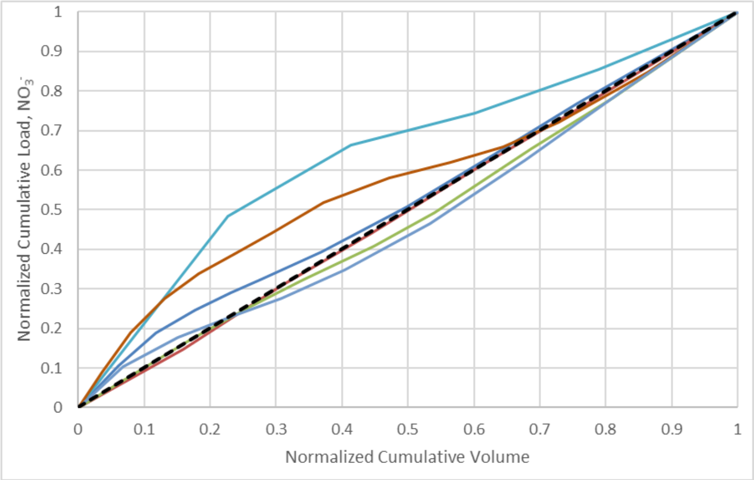


(c)



(d)

Figure 4 continued



(e)

Figure 4 continued

Table 4. Summary statistics of FF_{30}: p-values from Wilcoxon Rank Sum test and % events $FF_{30} > 0.3$

	$P_{30\%}$	Events with $FF_{30}>0.3$
Second Creek		
Fecal Coliform	0.5540	53%
E. coli	<u>0.0704</u>	65%
TSS	**< 0.0001**	88%
Cu^{2+}	**0.0006**	81%
NO_3^-	0.5192	41%
Third Creek		
Fecal Coliform	0.1724	71%
E. coli	0.1724	71%
TSS	**0.0493**	83%
Cu^{2+}	**0.0204**	86%
NO_3^-	0.6832	43%
Williams Creek		
Fecal Coliform	0.4002	63%
E. coli	0.4002	63%
TSS	<u>0.0821</u>	75%
Cu^{2+}	0.6823	57%
NO_3^-	0.9999	50%

p-value < 0.05 in bold
p-value < 0.1 in underline

confidence interval, respectively). It is important to note, this study expanded statistical significance to 0.1 to allow greater understanding of relationships within the dataset.

The first flush strength for the constituents at Second Creek, Third Creek, and Williams Creek, respectively, are as followed: TSS>Cu^{2+}>*E. coli*>NO_3^->fecal coliform, Cu^{2+}> TSS>fecal coliform> *E. coli*> NO_3^-, and TSS>*E. coli*>fecal coliform>Cu^{2+}>NO_3^- respectively. Even though FF_{30} was high for some pollutants, like Third Creek's fecal coliform (Table 3), no statistical significance was found. This is most likely due to small sample numbers or high variability in the data as evidenced by the relative standard deviation (RSD) of the pollutant. Hathaway and Hunt (2011) similarly performed a Wilcoxon signed-rank test on their FF_{30} data results from a 5.1 ha watershed in North Carolina and found TSS and fecal coliform to be significantly different from 0.3 ($p<0.05$).

Dissolved constituents, like Cu^{2+} and NO_3^-, typically do not show a consistent or strong first flush and this is the same conclusion for NO_3^- at every watershed (Characklis & Wiesner, 1997; Lee et al., 2002). However, Cu^{2+} is the second strongest constituent at Second Creek and the strongest at Third Creek. A study by Sansalone and Buchberger (1997) found a pronounced first flush for dissolved Cu from an urban roadway for lateral pavement sheet flow; although, it should be noted that this was determined by using a first flush definition of $m(t) > v(t)$. Using the first flush definition of 50% pollutant delivered in the first 25% volume, a different study found only 21% events for Cu^{2+} and 22% events for NO_3^- had a first flush from an all impervious roadway (Flint & Davis, 2007). The same study also found only 17% of events had a first flush for TSS. First flush presence of Cu^{2+} is more frequently identified in this study (Table 4), however it is lower in strength than the TSS first flush frequency. These differences in the strength of various pollutants between watersheds suggests that site specific variables influence pollutant transport trends.

Past studies have not found *E. coli* to have a consistent first flush with a mean FF_{30} typically around 0.3 and 0.4 in two notable studies (Hathaway & Hunt, 2011; D. McCarthy, 2009). While Second Creek and Williams Creek mean FF_{30} are 0.32 and 0.40, Third Creek's mean FF_{30} is 0.42, suggesting a more consistent first flush effect may be occurring at Third Creek. While Second Creek may have mean/median FF_{30} of 0.32, it was found to have a statistically significant first flush effect ($p < 0.1$). It is important to note that in this study and others, *E. coli* has highly variable FF_{30} results. This variability makes analyzing *E. coli* FF_{30} difficult, leading to a regression analyses later in this paper to determine possible explanatory variables. D. McCarthy (2009) did not find *E. coli* to have a consistent first flush for most catchments; but did indicate an end flush was occasionally occurring. Using the traditional method, end flushes were not frequently observed in this study.

TSS had the highest FF_{30} results (Table 3) exhibiting significantly higher values than 0.3 for all three watersheds ($p<0.1$). Past literature supports these findings, with some degree of sediment first flush generally being present and seemingly more perceptible than other pollutants (Taebi & Droste, 2004). As an example, studies such as Li et al. (2007) found TSS FF_{30} event values between 0.52 and 0.72, while Taebi and Droste (2004) found 30.9% of TSS load to be delivered in the first 20% of event volume.

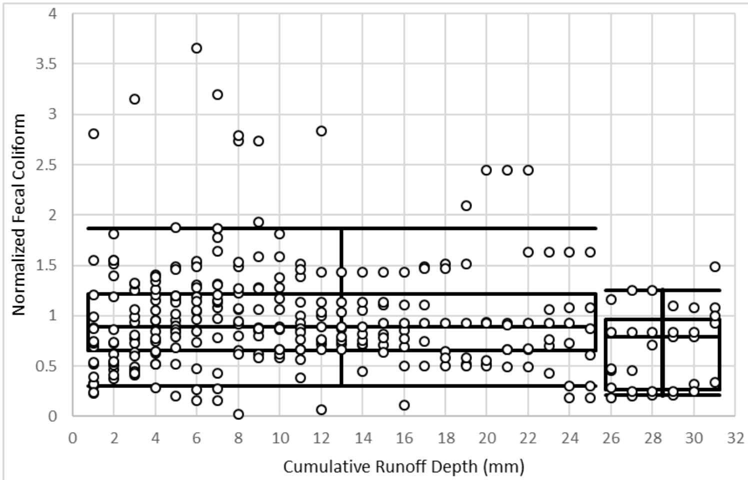
Overall, multiple studies concluded some sort of flush occurred for TSS (Bach et al., 2010; Bertrand-Krajewski et al., 1998; Deletic, 1998; Flint & Davis, 2007; Hathaway & Hunt, 2011; Sansalone & Cristina, 2004).

Using this method, a first flush is observed in many pollutants and appears to be present in the majority (>75%) of events for TSS (all sites), and Cu^{2+} (Second Creek and Third Creek). While past studies have found TSS to have consistent first flushes, they are weaker than the FF_{30} values found in this study. This may be due to the site locations, where open channel streams are able to contribute resuspended sediment to the flow during storm events. Therefore, in-stream processes are most likely influencing higher FF_{30} results. Due to high velocities in-stream during a storm event, observing a first flush of TSS appears reasonable. Streambeds are also home to microbes as past literature has explained (Muirhead et al., 2004; Surbeck et al., 2016). This would explain *E. coli's* first flush significance at Second Creek, and *E. coli's* first flush event frequency at Third Creek and Williams Creek.

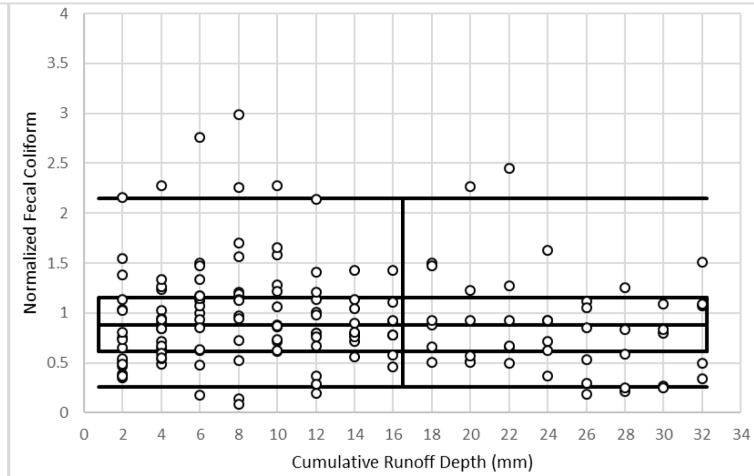
3.2 Bach Slice Method: Modified First Flush Analysis

As previously discussed, the traditional first flush analysis method does not consider storm event size which may result in skewed conclusions, that is, larger storms may wash off more pollutants resulting in source limited conditions as surface stores are depleted. Thus, a new methodology was investigated for comparison. Slice size selection was directed by two things: (1) Bach et al. (2010) stated slice size sensitivity was not significant when comparing sizes between 0.5 and 3 mm and (2) they suggested the slice size should be chosen as the smallest slice that has at least one explicitly measured data point rather than all data within that slice containing interpolated data. A slice size of 1 mm met that criteria; however, the results of the Wilcoxon Rank Sum Test used to evaluate significant differences in slices showed only significant differences in the latter part of the cumulative runoff depth. For example, a new slice group started after an initial slice size of 28 mm for Second Creek's TSS. This outcome was not consistent with other studies that have noted a first flush effect for TSS, and the traditional first flush analysis implied a strong flushing effect which prompted further analysis. To verify the slice size sensitivity for these data, calculations were rerun with a 2 mm slice size, and results showed contaminants that typically have a first flush (i.e. TSS) still were not being represented as such. Next, instead of using calculations with the 2 mm slice size, data were combined by grouping data from 1 mm and 2 mm renamed 1-2, 3 mm and 4 mm renamed 3-4 and so on. This method seemed to more accurately represent the data and various pollutants' first flush. This is likely due to increased statistical power gained with a larger number of data points in each slice (Figure 5). Final results are shown in Figure 6-8 using the 2 mm grouping method.

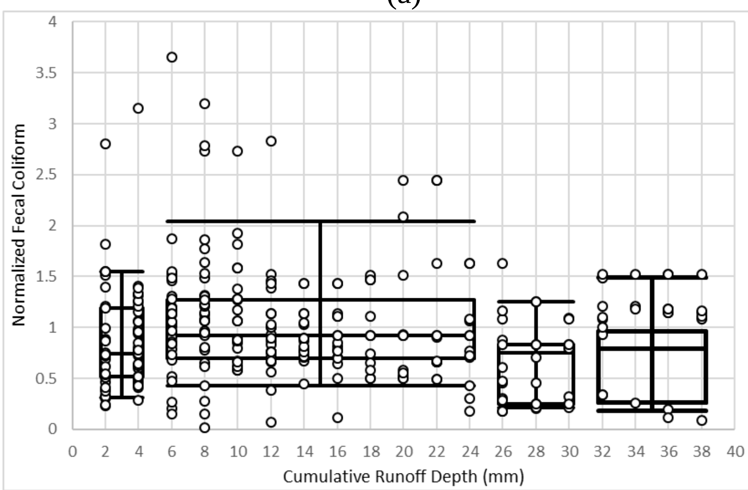
Second Creek had the highest maximum runoff depth due to a number of larger events being monitored relative to the other two sites. A first flush occurred when two or more box and whisker plots, or groups, were present. The only pollutants that did not have more than one group are *E. coli* (Fig. 7b) and NO_3^- (Fig. 7e) from Third Creek. Also of note were Second Creek's fecal coliform (Fig. 6a), *E. coli* (Fig. 6b), and NO_3^- (Fig.



(a)



(b)



(c)

Figure 5. Modified first flush slice size methods (a) 1 mm slice size (b) 2 mm slice size and (c) combined data from 1 mm slices into 2 mm slice size intervals for fecal coliform at Second Creek

Figure 6. Modified first flush grouped Box and Whisker Plots of normalized pollutant concentrations vs cumulative runoff depths for Second Creek site. (a) fecal coliform, (b) *E. coli*, (c) TSS, (d) Cu^{2+}, and (e) NO_3^-

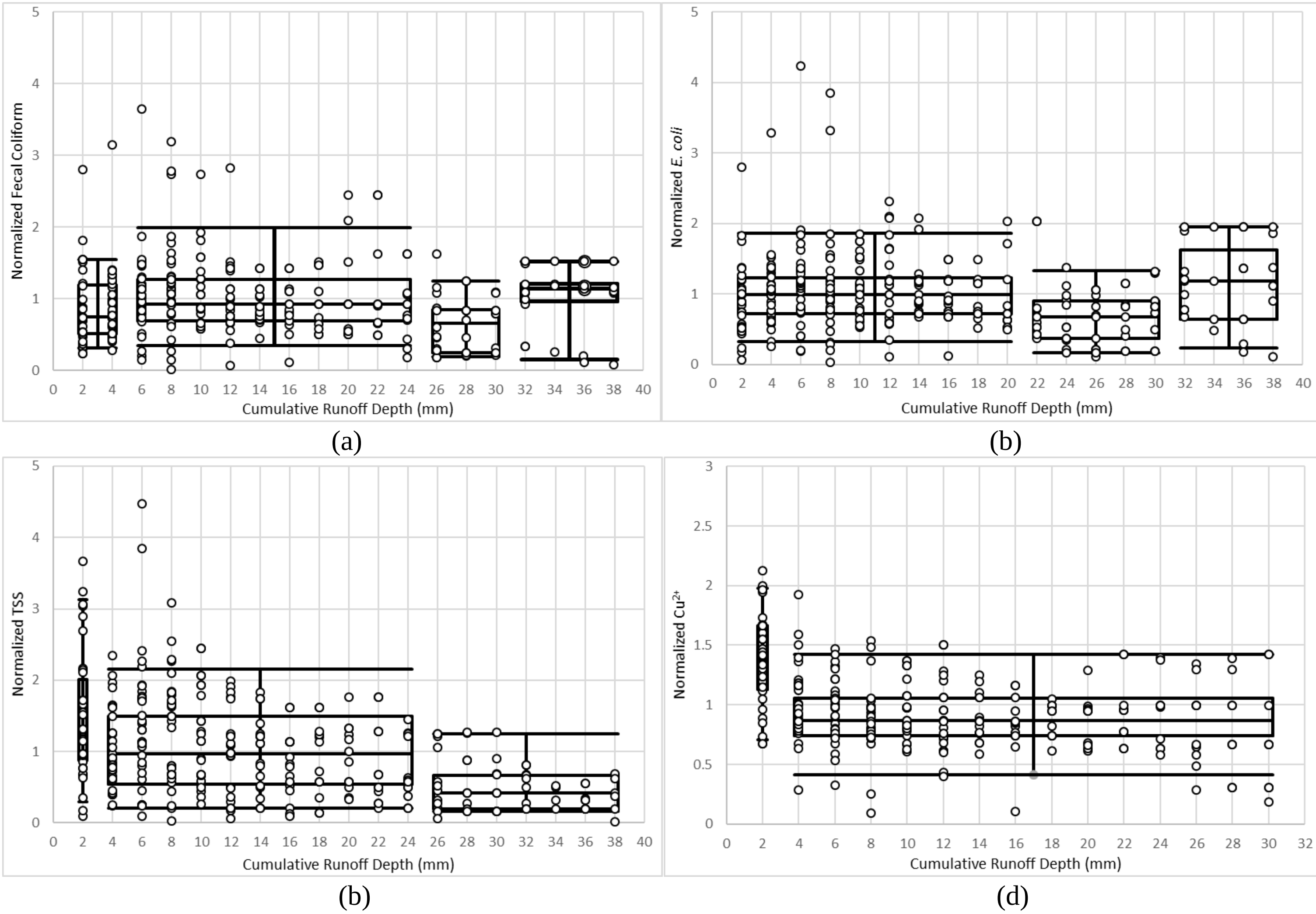
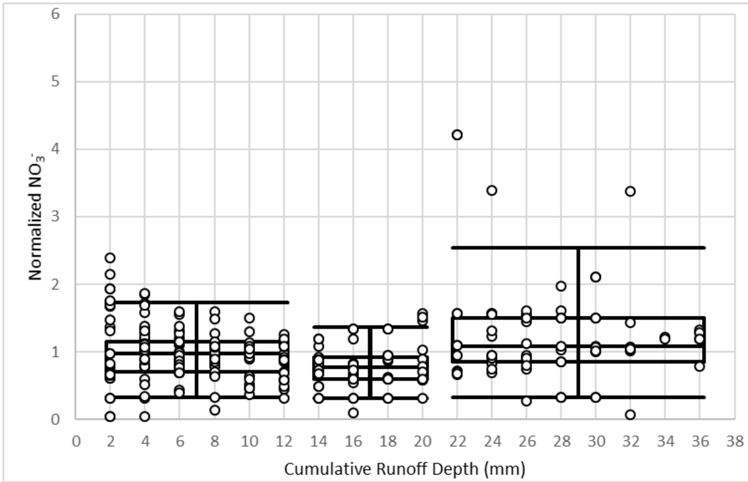


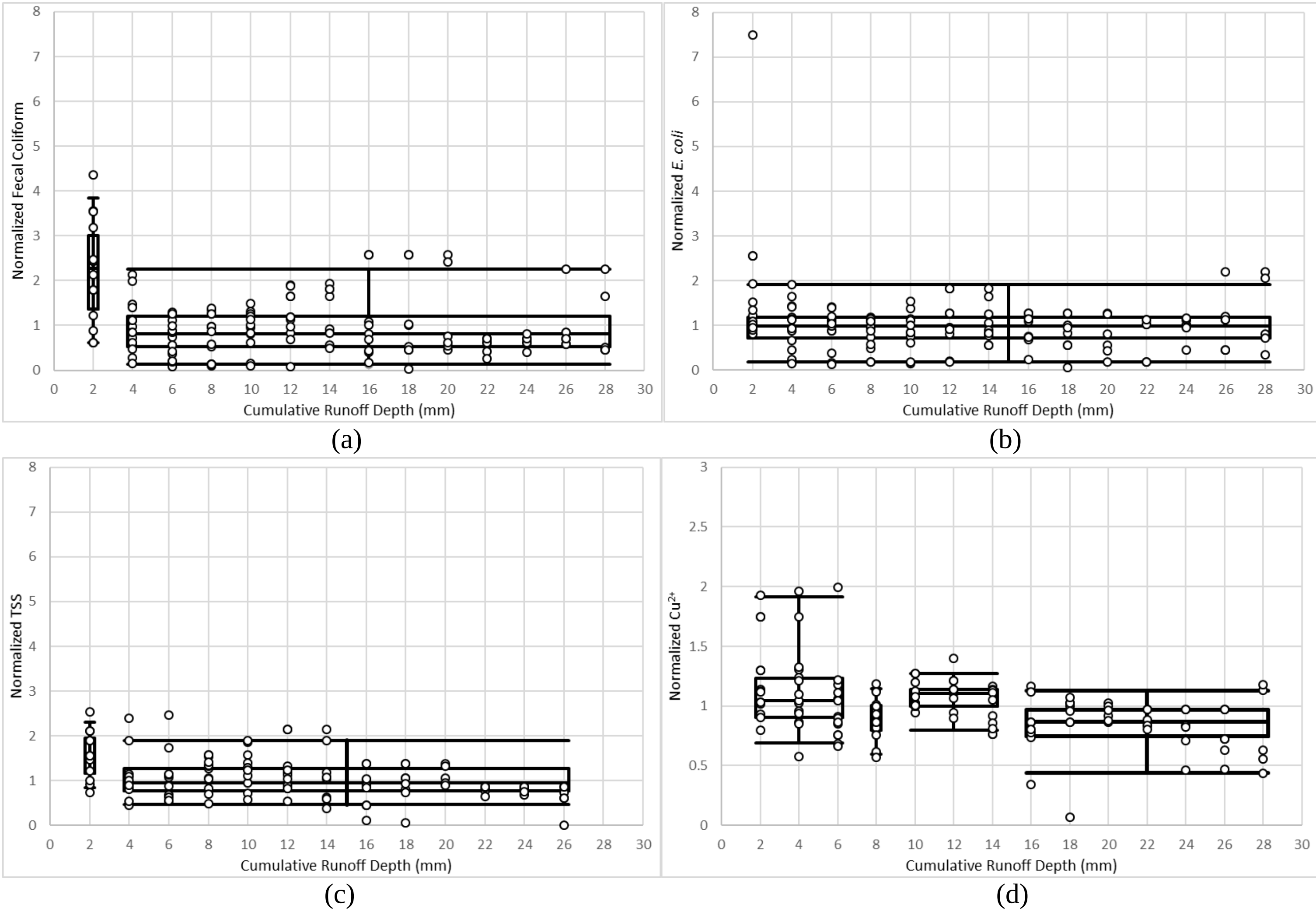
Figure 6 continued



(e)

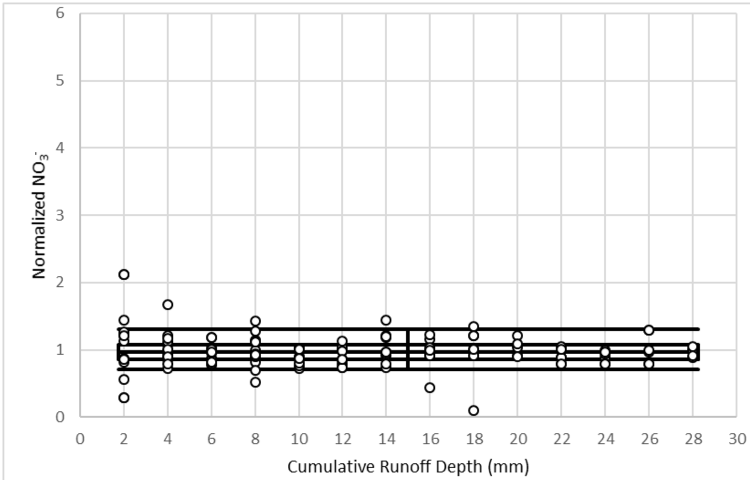
Figure 6 continued

Figure 7. Modified first flush grouped Box and Whisker Plots of normalized pollutant concentrations vs cumulative runoff depths for Third Creek site. (a) fecal coliform, (b)*E. coli*, (c) TSS, (d) Cu^{2+}, and (e) NO_3^-



(a) (b) (c) (d)

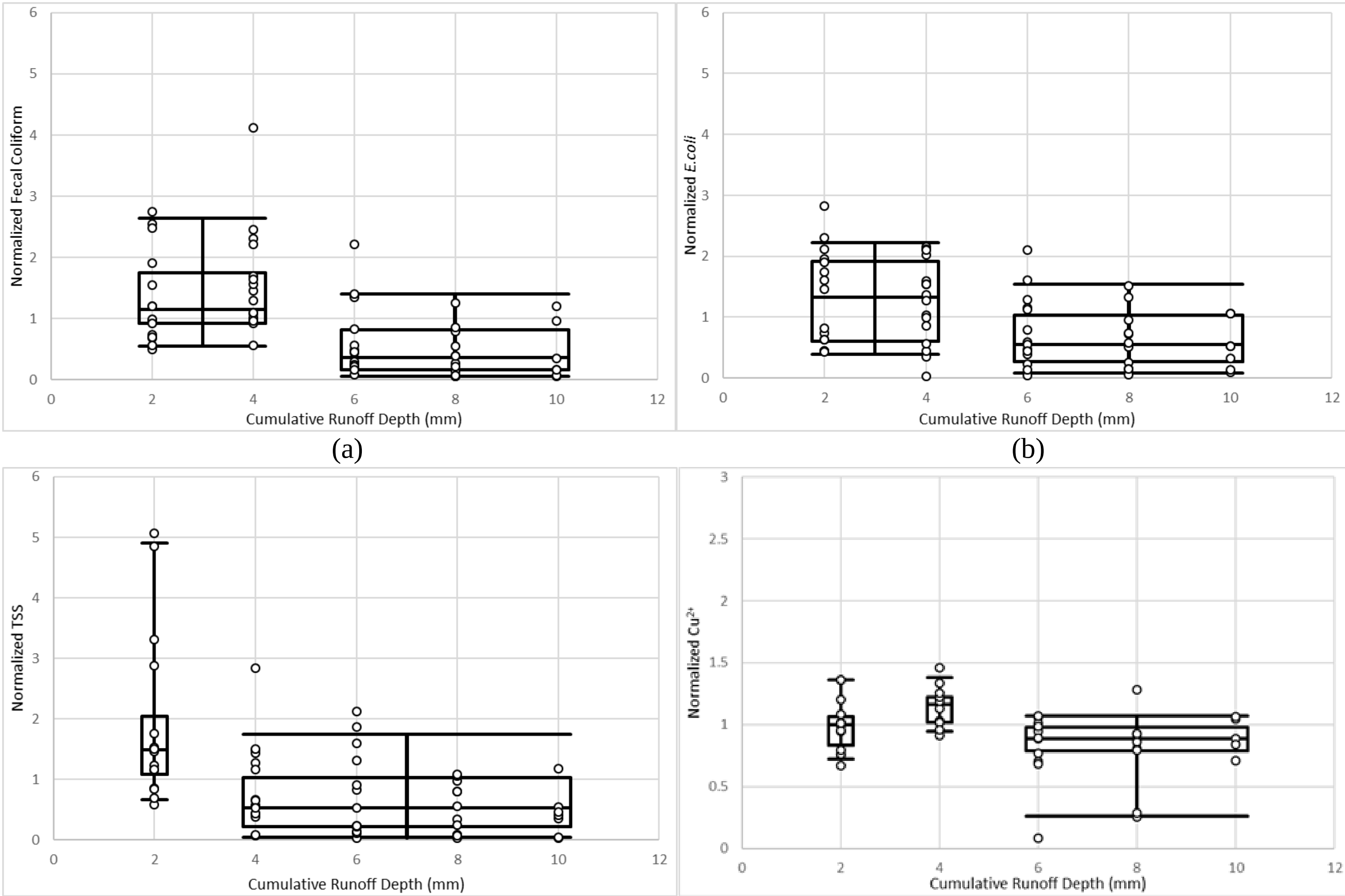
Figure 7 continued



(e)

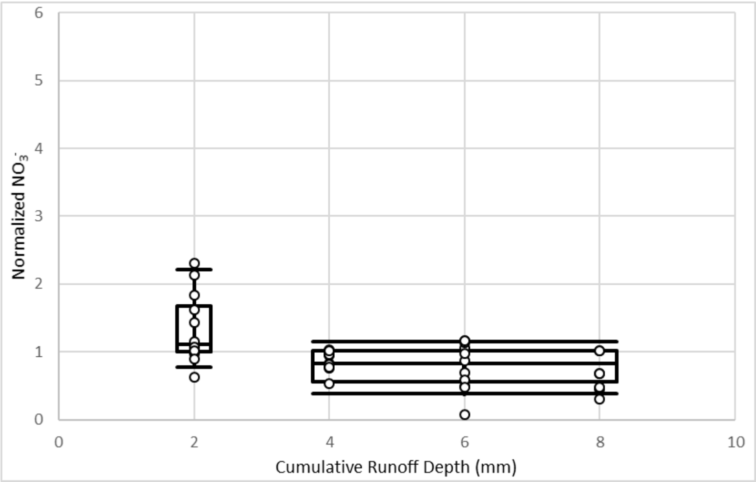
Figure 7 continued

Figure 8. Modified first flush grouped Box and Whisker Plots of normalized pollutant concentrations vs cumulative runoff depths for Williams Creek site. (a) fecal coliform, (b) *E. coli*, (c) TSS, (d) Cu^{2+}, and (e) NO_3^-



(a) (b) (c) (d)

Figure 8 continued



(e)

Figure 8 continued

6e); Third Creek's Cu^{2+} (Fig. 7d); and Williams Creek's Cu^{2+} (Fig. 8d) boxplots where inconsistent trends were observed, that is, both increases and decreases were noted in groups as runoff depth increased. The other two studies that used this method did not observe such trends but also did not have more than 3 slices or utilize the same procedure for grouping slices.

Using the modified method, the first flush strength and first flush depth can be quantified. The strength of the first flush, or P_{FF}, is defined as the p-value of the Wilcoxon Rank Sum analysis between the first and last group. However, this method assumes that the groups decline in concentration with increased runoff. In this study, groups do not decline consistently thus this may not be an accurate way to determine the strength of the first flush using the modified method. The modified method also determines first flush volume, V_{FF}, to be the runoff depth up to the beginning of the last slice. If a pollutant has a high V_{FF}, this indicates the high concentration in the first flush takes a significant amount of runoff to be reduced. Results from this study using the modified first flush strength definition are presented in Table 5.

The majority of contaminants exhibited a significant first flush, including: Second Creek's fecal coliform, TSS, and Cu^{2+}; Third Creek's fecal coliform, TSS, Cu^{2+}; and all pollutants from Williams Creek. Williams Creek, the smallest watershed, exhibited a first flush for all pollutants. As discussed previously, this could be due to the smaller watershed. However, it could be due to the limited runoff depth represented because of limited data. Second Creek's fecal coliform is the only significant first flush with a

Table 5. First flush strength, P_{FF}, for modified first flush method for pollutants at each watershed.

	P_{FF}	V_{FF} (mm)
Second Creek		
Fecal Coliform	0.0846	30
E. coli	0.3159	30
TSS	**< 0.0001**	24
Cu^{2+}	**< 0.0001**	2
NO_3^-	0.2908	20
Third Creek		
Fecal Coliform	**< 0.0001**	2
E. coli	0.1350*	2*
TSS	**0.0016**	2
Cu^{2+}	**< 0.0001**	14
NO_3^-	0.2124*	8*
Williams Creek		
Fecal Coliform	**< 0.0001**	4
E. coli	**0.0009**	4
TSS	**< 0.0001**	2
Cu^{2+}	**0.0077**	4
NO_3^-	**0.0233**	2

*Results did not exhibit first flush (1 box), lowest p-value between slices reported
p-value < 0.05 in bold
p-value < 0.1 in underline

p-value significance of $p<0.1$. TSS and Cu^{2+} are the only pollutants to show a first flush at all three sites when $p<0.05$.

For those pollutants that did not exhibit a first flush or for which the first flush was only noted for one group, the lowest p-value and the runoff depth at which a first flush is likely to occur is reported. From Table 5, aside from the two pollutants that did not have more than one group, Second Creek's fecal coliform, *E. coli*, and NO_3^- (Fig. 6a, 6b, 6e) do not have a statistically significant ($p<0.05$) first flush between the first and last slices and are also considered to exhibit one of the observed inconsistent trends. This is most likely due to the observed end flush occurring on the last slice for both contaminates. This is not the first time an end flush has been seen in literature, particularly for indicator bacteria. Other studies suggested end flushes could be due to wastewater intrusions or watershed land use characteristics (Hathaway & Hunt, 2011; D. McCarthy, 2009). For example, Hathaway and Hunt (2011) had some events with end flushes from a catchment containing high pervious areas like residential yards, where domestic animals would contribute to concentrations of indicator bacteria. (Bach et al., 2010) noted the modified method was not designed to quantify end flushes in its analysis, but believes it can easily be adjusted for this purpose. However, the modified method does allow for end flushes to be observed from slice groupings, without quantification. The other two graphs with inconsistent trends, Cu^{2+} from Third Creek and Williams Creek are found to be statistically significant and this is most likely due to the last slice being less than the first slice/box. Therefore, the graphs with inconsistent trends still can show a significant first flush.

3.3 Synthesis of Method Outcomes

The traditional method and the modified method were synthesized to understand how the results compared. When looking at median FF_{30} from the traditional method, all of the pollutants with inconsistent trends or one group have a median between 0.29-0.33. Whereas, Second Creek's TSS and Cu^{2+}; Third Creek's fecal coliform and TSS; and Williams Creek's fecal coliform, *E. coli*, and TSS have FF_{30} medians between 0.36-0.47, which is typically corroborated by a first flush in the boxplot graph generated as part of the Bach et al. (2010) method. This relationship suggests FF_{30} medians are typically higher for those boxplot graphs expressing a first flush; that is, a first flush without inconsistent trends.

Simple observation of the graphs (Figures 6-8) also suggested the first flush strength when compared to the FF_{30}. These first flush pollutants plots also have a first grouping which constitutes 2-4 mm of runoff depth followed by other groups with decreasing values. The smaller the runoff volume represented by the first group, and the higher the magnitude of this group in comparison to subsequent groups, the higher the FF_{30} median. The P_{FF} values support this idea. For example, Williams Creek NO_3^- (Figure 8e) has two groupings, which has a 2 mm initial slice depth with the second one decreasing and a FF_{30} median of 0.32. The P_{FF} for Williams Creek NO_3^- is 0.0233 which is slightly weaker than Williams Creek TSS P_{FF} of <0.0001 which also has a 2 mm initial slice depth, but higher group magnitude. Thus, there is a first flush for Williams Creek NO_3^- , but it is not as strong as Williams Creek TSS, for example. Conversely, Williams

Creek fecal coliform has the same P_{FF} strength as Williams Creek TSS with a 4 mm initial slice depth, and a higher FF_{30} median of 0.36. This indicates considering initial slice depth size and FF_{30} may play a factor in determining first flush strength, but must be accompanied by P_{FF} to accurately provide first flush quantification.

Table 6 shows a summary of each method's outcomes. The traditional method also determined the TSS from each watershed and Cu^{2+} from Second Creek and Third Creek to have a significant first flush, but failed to recognize the other pollutants the modified method identified. Additionally, the traditional method determined Second Creek's *E. coli* to have a significant first flush which was not identified in the modified first flush strength method.

3.4 Influence of Antecedent Climate and Event Specific Parameters

To better understand the causes of the variable FF_{30} values between storms, the Spearman Rank test was run on each pollutant for Second Creek using various explanatory variables and FF_{30} values. There were not enough events for the sample results to be valid for Third Creek and Williams Creek, thus they are not discussed herein. Table 7 shows the spearman rank coefficient (r_s) and its p-value ($p<0.1$ and $p<0.05$ in bold) for each pollutant correlation to antecedent climate, storm, and runoff variables.

Table 6. Summary of method analyses

		$P_{30\%}$	P_{FF}	Median FF_{30}	Initial Group Size (mm)	V_{FF} (mm)	Boxplot Trend
Second Creek	Fecal Coliform	0.554	0.0846	0.31	4	30	Inconsistent
	E. coli	0.0704	0.3159	0.32	20	30	Inconsistent
	TSS	**< 0.0001**	**< 0.0001**	0.46	2	24	First flush
	Cu^{2+}	**0.0006**	**< 0.0001**	0.37	2	2	First flush
	NO_3^-	0.5192	0.2908	0.30	12	20	Inconsistent
Third Creek	Fecal Coliform	0.1724	**< 0.0001**	0.46	2	2	First flush
	E. coli	0.1724	0.1350*	0.33	28	2*	No first flush
	TSS	**0.0493**	**0.0016**	0.39	2	2	First flush
	Cu^{2+}	**0.0204**	**< 0.0001**	0.32	6	14	Inconsistent
	NO_3^-	0.6832	0.2124*	0.29	28	8*	No first flush
Williams Creek	Fecal Coliform	0.4002	**< 0.0001**	0.36	4	4	First flush
	E. coli	0.4002	**0.0009**	0.38	4	4	First flush
	TSS	0.0821	**< 0.0001**	0.47	2	2	First flush
	Cu^{2+}	0.6823	**0.0077**	0.30	2	4	Inconsistent
	NO_3^-	0.9999	**0.0233**	0.32	2	2	First flush

*Results did not exhibit first flush (1 box), lowest p-value between slices reported
p-value < 0.05 in bold
p-value < 0.1 in underline

Table 7. Spearman Rank correlation analysis between flow, rainfall, and antecedent climate explanatory variables and pollutants

		Fecal Coliform		*E.coli*		TSS		Cu^{2+}		NO_3^-	
		r_s	p-value	r_s	p-value	r_s	p-value	r_s	p-value	r_s	p-value
Antecedent Climate	Max. $Temp_{28\ days}$	-0.4601	0.0631	-0.5411	**0.0249**					-0.5041	**0.0456**
	Max. $Temp_{2\ days}$	-0.4457	0.0730	-0.5832	**0.0140**					-0.4726	0.0645
	Min. $Temp_{28\ days}$	-0.4601	0.0631	-0.5411	**0.0249**					-0.5041	**0.0456**
	Min. $Temp_{2\ days}$	-0.4416	0.0759	-0.5878	**0.0313**					-0.4848	0.0570
	Min. $RH_{28\ days}$							-0.4801	0.0598	0.5395	**0.0310**
	Total $Rainfall_{28\ days}$					0.8671	**0.0003**				
	$ADWD_{0.1}$			0.4611	0.0745						
Rainfall	Total Rainfall							-0.6014	**0.0297**	0.4931	0.0869
	Ave. Rainfall Intensity							-0.4993	0.0824	-0.7036	**0.0073**
	Max. Rainfall Intensity	-0.4580	0.0996					-0.7095	**0.0066**	-0.5384	0.0577
Storm-water flow	Max. Flow Rate	-0.6233	**0.0075**	-0.4638	0.0608			-0.6145	**0.0113**		
	Ave. Flow Rate	-0.5080	**0.0374**					-0.6041	**0.0132**		

3.4.i Antecedent Climate

Antecedent temperature was found to negatively correlate to fecal coliform, *E. coli*, and NO_3^- FF_{30}. Similarly, Hathaway and Hunt (2011) found fecal coliform was negatively correlated to antecedent temperature; however, no correlation was found between temperature and *E. coli*. In contrast, McCarthy (2009) found *E. coli* to have a positive correlation with temperature, and concluded this was due to *E. coli* die-off associated with increased temperature. McCarthy (2009) hypothesized that this lower level of *E. coli* would be depleted in the beginning of the event resulting in a strong first flush. However, this study found *E. coli* to negatively correlate to temperature, meaning with higher temperatures in days leading up to event, a weaker first flush would occur for *E. coli*. Hathaway and Hunt (2011) found seasonal differences were significant for both fecal coliform and *E. coli*. The study concluded microbes were limited during the winter months, supporting this study's negative correlation between temperature and microbes. The first flush of microbes during summer months were weak because there are most likely higher concentrations of microbes versus in the winter months when the microbes are more limited. *E. coli* FF_{30} is the only constituent to correlate to ADWP (antecedent dry weather period). As ADWP decreased, the FF_{30} decreased. This trend is likely explained by flushing within the system, that is, if frequent rainfall has been occurring, the microbial store is depleted resulting in more consistent concentrations. While correlations vary in magnitude between indicator bacteria, antecedent conditions, like temperature, do appear to have an impact on microbial export patterns. Antecedent conditions have been observed as explanations for indicator bacteria variability and behavior (Hathaway et al., 2010).

The explanation for the negative correlation between temperature and NO_3^- is uncertain. This could be due to source limited supply of fertilizers. In winter months, fertilizers, the potential source of NO_3^-, are not as frequently used than it is in the summer months. Therefore, there is a better chance the NO_3^- could be washed off from storm events resulting in a stronger first flush. However, the correlation could be due to another variable that is related to temperature. One study looking at nutrient variability in stormwater runoff found it difficult to correlate to variables except for catchment area. The lack of data regarding homeowners use and timing of fertilizer most likely played an important role in the nutrients variability (Toran & Grandstaff, 2007). This missing data may have some correlation to low temperatures or other explanatory variables and its correlation to NO_3^-.

The antecedent total rainfall was positively and strongly correlated to TSS. There were no other variables that correlated to TSS despite some previous literature which showed the TSS first flush related to other rainfall related parameters such as peak rainfall intensity, storm duration, antecedent dry weather period Gupta and Saul (1996) Lee et al. (2002) Li et al. (2007) Deletic (1998) and Taebi and Droste (2004). In this study, TSS is positively and strongly correlated to the 28-day antecedent total rainfall depth, meaning that monitored rain events following a high magnitude of rainfall in the previous 28 days were more likely to have a high first flush. This is logical, as antecedent rains likely wash away sediment and leave less in the surface store for subsequent events. The reason no other explanatory variables correlated to TSS FF_{30}, despite that trend being recognized in other studies, could be because there is a response for sediment no matter climate, rainfall, or flow factors. Specifically, this study being performed in an open channel is a significant factor. Sediments in-stream are most likely suspended on some level no matter the flow or rainfall intensities.

3.4.ii Rainfall

Additionally, three rainfall variables, total rainfall, average rainfall intensity, and max rainfall intensity, correlated to the dissolved pollutants, Cu^{2+} and NO_3^-. As shown through past research, a first flush for Cu^{2+} is seen more frequently from roadways, and could be the source of Cu^{2+} for this study. Therefore, for smaller events, most of the runoff is most likely delivered from connected impervious surfaces. From the correlation analysis, smaller events have stronger first flushes. Whereas, larger storm events have other pervious areas contributing, possibly diluting Cu^{2+} concentrations, resulting in weaker first flushes for Cu^{2+}. However, it could also be the opposite. Larger events could be contributing more source of Cu^{2+}, resulting in a weak first flush, and small events have limited sources of Cu^{2+} from connected impervious surfaces. Based on the previous literature regarding first flush of Cu^{2+}, it would seem the first reasoning is most likely occurring.

3.4.iii Stormwater Flow

Microbes were negatively correlated to flow rate, so with high flow rates a weak first flush was exhibited. Surbeck, Jiang, Ahn, and Grant (2006) found microbes to increase abruptly and remain elevated after flow increase in-stream from three storm events monitored at three different locations in southern California. Therefore, higher flows could result in elevated levels of microbes that stay suspended over an event, resulting in a weak first flush. Thus, higher flow rates mean more microbes are available for flushing. A stronger first flush for microbes may occur at lower flow rates because not as many microbes were mobilized.

Flow rates were also negatively correlated to Cu^{2+}. Characklis and Wiesner (1997) found dissolved constituents typically followed flow patterns more closely than sediment particles which supports this study's findings. When there are higher flow rates, the dissolved constituent, Cu^{2+}, was more available resulting in a weak first flush. However, this does not explain why no correlation was found between flow rates and NO_3^-.

4 CONCLUSION

Three urban streams were monitored for 32 storm total events for fecal coliform, *E. coli*, TSS, Cu^{2+}, and NO_3^-. Two different methods were utilized to determine the first flush for each constituent. One traditionally used in literature that defined the strength of the first flush as the fractional mass delivered in the first 30% event volume. And the other proposed by Bach et al. (2010), that takes account of storm size, and quantifies first flush volume and first flush strength. Results from the traditional method show Second Creek's *E. coli*, TSS, and Cu^{2+}; Third Creek's TSS and Cu^{2+}; and Williams Creek TSS median FF_{30} were statistically different than 30% with a p-value <0.05 or <0.1. When compared to past research which utilized the traditional method, this study seems to have more consistent and significant first flush effects for more pollutants. It could be concluded this is due to in-stream influences, as many other studies monitored stormwater runoff prior to its entrance into open channel streams. A significant first flush from the modified method occurred for Second Creek's fecal coliform, TSS, and Cu^{2+}; Third Creek's fecal coliform, TSS, and Cu^{2+}; and all pollutants from Williams Creek. Thus, the modified method identified more pollutants that exhibited a first flush than the traditional method results. Additionally, the modified method had three factors, P_{FF}, V_{FF}, and initial group size, that further quantified first flush characteristics. However, after comparing the two methods, it appears the FF_{30} median from the traditional method supports conclusions from the modified method.

Ultimately, this study brings insight as to how in-stream processes affect pollutant load. Past studies have shown the first flush for sediment and microbes do occur in-stream from increased flows (Muirhead et al., 2004; Surbeck et al., 2016). This study found a significant first flush occurred for TSS in-stream at all site locations. While fewer sites in this study found a significant first flush for microbes compared to TSS, a consistent first flush trend for fecal coliform at two locations and E. coli at one location were observed based on the modified method. Microbe first flush strength results between locations are most likely due to specific site characteristics and microbe variability. Previous studies have found a weaker first flush and less frequently identified for TSS and microbes than this study; however, this is most likely due to their stormwater outfall collection sites. Additionally, some microbes appeared to exhibit end flush trends which were trends also reported by previous literature. The modified method also concluded Cu^{2+} to be the second strongest pollutant with a significant first flush at each site location. There has not been another study that has found similar first flush strength for Cu^{2+}. Therefore, it could be due to in-stream influences or a watershed characteristic that could be at each study location.

Additionally, this study contributes to the studies who have used the modified method by Bach et al. (2010). This study showed more complex results than previous studies that have utilized this method. For example, this study has more than 3 groups present on a graph with some displayed less consistent trends than previous identified. (Bach et al., 2010) and (Hathaway et al., 2016) did not have graphs with more than 3 boxes or show obvious end flushes. This could be due to in-stream collection; however, it cannot be said with confidence until further studies have utilized the modified method.

From the correlation analysis between the first flush and possible explanatory variables, microbes appeared to be predominately influenced by antecedent climate, Cu^{2+} and NO_3^- appeared to mostly correlate with rainfall factors, and flow rate variables correlated to microbes and Cu^{2+} first flush strength. However, NO_3^- first flush strength was also heavily influenced by antecedent parameters. Only variable to correlate to the first flush strength of TSS was ADWP, and this could be because sediment is resuspended in-stream no matter rainfall or stormflow effects.

Further studies, especially in-stream, looking into the relationships between water quality parameters and their influences could enhance water quality model performance. Results from this study suggest the first flush is significant when collected in-stream, and is more significant than other studies that collected from stormwater outfall. Therefore, there is a need to incorporate an in-stream pollutant model with current water quality models. Further analysis and studies regarding pollutants and their influences would enable the possibility for these upgraded models.

LIST OF REFERENCES

Bach, P. M., McCarthy, D. T., & Deletic, A. (2010). Redefining the stormwater first flush phenomenon. *Water Research, 44*(8), 2487-2498. doi:10.1016/j.watres.2010.01.022

Batroney, T., Wadzuk, B. M., & Traver, R. G. (2010). Parking Deck's First Flush. *Journal of Hydrologic Engineering, 15*(2), 123-128. doi:10.1061/(asce)he.1943-5584.0000167

Bertrand-Krajewski, J. L., Chebbo, G., & Saget, A. (1998). Distribution of pollutant mass vs volume in stormwater discharges and the first flush phenomenon. *Water Research, 32*(8), 2341-2356. doi:10.1016/s0043-1354(97)00420-x

Characklis, G. W., Dilts, M. J., Simmons, O. D., Likirdopulos, C. A., Krometis, L. A. H., & Sobsey, M. D. (2005). Microbial partitioning to settleable particles in stormwater. *Water Research, 39*(9), 1773-1782. doi:10.1016/j.watres.2005.03.004

Characklis, G. W., & Wiesner, M. R. (1997). Particles, metals, and water quality in runoff from large urban watershed. *Journal of Environmental Engineering-Asce, 123*(8), 753-759. doi:10.1061/(asce)0733-9372(1997)123:8(753)

Deletic, A. (1998). The first flush load of urban surface runoff. *Water Research, 32*(8), 2462-2470. doi:10.1016/s0043-1354(97)00470-3

Dotto, C. B. S., Kleidorfer, M., Deletic, A., Fletcher, T. D., McCarthy, D. T., & Rauch, W. (2010). Stormwater quality models: performance and sensitivity analysis. *Water Science and Technology, 62*(4), 837-843. doi:10.2166/wst.2010.325

Flint, K. R., & Davis, A. P. (2007). Pollutant mass flushing characterization of highway stormwater runoff from an ultra-urban area. *Journal of Environmental*

Engineering-Asce, 133(6), 616-626. doi:10.1061/(asce)0733-9372(2007)133:6(616)

Geiger, W. (1984). *Characteristics of combined sewer runoff.* Paper presented at the Proceeding de la 3ème conférence internationale «Urban Storm Drainage», Göteborg.

Grable, J. L., & Harden, C. P. (2006). Geomorphic response of an Appalachian Valley and Ridge stream to urbanization. *Earth Surface Processes and Landforms, 31*(13), 1707-1720. doi:10.1002/esp.1433

Gupta, K., & Saul, A. J. (1996). Specific relationships for the first flush load in combined sewer flows. *Water Research, 30*(5), 1244-1252.

Hathaway, J. M., & Hunt, W. F. (2011). Evaluation of First Flush for Indicator Bacteria and Total Suspended Solids in Urban Stormwater Runoff. *Water Air and Soil Pollution, 217*(1-4), 135-147. doi:10.1007/s11270-010-0574-y

Hathaway, J. M., Hunt, W. F., & Simmons, O. D. (2010). Statistical Evaluation of Factors Affecting Indicator Bacteria in Urban Storm-Water Runoff. *Journal of Environmental Engineering-Asce, 136*(12), 1360-1368. doi:10.1061/(asce)ee.1943-7870.0000278

Hathaway, J. M., Tucker, R. S., Spooner, J. M., & Hunt, W. F. (2012). A Traditional Analysis of the First Flush Effect for Nutrients in Stormwater Runoff from Two Small Urban Catchments. *Water Air and Soil Pollution, 223*(9), 5903-5915. doi:10.1007/s11270-012-1327-x

Hathaway, J. M., Winston, R. J., Brown, R. A., Hunt, W. F., & McCarthy, D. T. (2016). Temperature dynamics of stormwater runoff in Australia and the USA. *Science of the Total Environment, 559*, 141-150. doi:http://dx.doi.org/10.1016/j.scitotenv.2016.03.155

Krometis, L. A. H., Characklis, G. W., Simmons, O. D., Dilts, M. J., Likirdopulos, C. A., & Sobsey, M. D. (2007). Intra-storm variability in microbial partitioning and microbial loading rates. *Water Research, 41*(2), 506-516. doi:10.1016/j.watres.2006.09.029

Lee, J. H., & Bang, K. W. (2000). Characterization of urban stormwater runoff. *Water Research, 34*(6), 1773-1780. doi:10.1016/s0043-1354(99)00325-5

Lee, J. H., Bang, K. W., Ketchum, L. H., Choe, J. S., & Yu, M. J. (2002). First flush analysis of urban storm runoff. *Science of the Total Environment, 293*(1-3), 163-175. doi:10.1016/s0048-9697(02)00006-2

Lee, J. H., Yu, M. J., Bang, K. W., & Choe, J. S. (2003). Evaluation of the methods for first flush analysis in urban watersheds. *Water Science and Technology, 48*(10), 167-176.

Li, L.-Q., Yin, C.-Q., He, Q.-C., & Kong, L.-L. (2007). First flush of storm runoff pollution from an urban catchment in China. *Journal of Environmental Sciences, 19*(3), 295-299.

McCarthy, D. (2009). A traditional first flush assessment of E. coli in urban stormwater runoff. *Water Science and Technology, 60*(11), 2749.

McCarthy, D. T., Hathaway, J. M., Hunt, W. F., & Deletic, A. (2012). Intra-event variability of Escherichia coli and total suspended solids in urban stormwater runoff. *Water Research, 46*(20), 6661-6670. doi:10.1016/j.watres.2012.01.006

Muirhead, R., Davies-Colley, R., Donnison, A., & Nagels, J. (2004). Faecal bacteria yields in artificial flood events: quantifying in-stream stores. *Water Research, 38*(5), 1215-1224.

Pachepsky, Y., Guber, A., Shelton, D., & Hill, R. (2009). *E. coli resuspension during an artificial high-flow event in a small first-order creek.* Paper presented at the EGU General Assembly Conference Abstracts.

Saget, A., Chebbo, G., & Bertrand-Krajewski, J.-L. (1996). The first flush in sewer systems. *Water Science and Technology, 33*(9), 101-108.

Sansalone, J. J., & Buchberger, S. G. (1997). Partitioning and first flush of metals in urban roadway storm water. *Journal of Environmental Engineering-Asce, 123*(2), 134-143. doi:10.1061/(asce)0733-9372(1997)123:2(134)

Sansalone, J. J., & Cristina, C. M. (2004). First flush concepts for suspended and dissolved solids in small impervious watersheds. *Journal of Environmental Engineering-Asce, 130*(11), 1301-1314. doi:10.1061/(asce)0733-9372(2004)130:11(1301)

Sansalone, J. J., Koran, J. M., Smithson, J. A., & Buchberger, S. G. (1998). Physical characteristics of urban roadway solids transported during rain events. *Journal of Environmental Engineering-Asce, 124*(5), 427-440. doi:10.1061/(asce)0733-9372(1998)124:5(427)

Stein, E. D., Tiefenthaler, L. L., & Schiff, K. C. (2007). Sources, patterns and mechanisms of storm water pollutant loading from watersheds and land uses of the greater Los Angeles area, California, USA. *Southern California Coastal Water Research Project. Technical Report, 510*.

Surbeck, C. Q., Jiang, S. C., Ahn, J. H., & Grant, S. B. (2006). Flow fingerprinting fecal pollution and suspended solids in stormwater runoff from an urban coastal watershed. *Environmental Science & Technology, 40*(14), 4435-4441. doi:10.1021/es060701h

Surbeck, C. Q., Shields, F. D., & Cooper, A. M. (2016). Fecal Indicator Bacteria Entrainment from Streambed to Water Column: Transport by Unsteady Flow over a Sand Bed. *Journal of Environmental Quality, 45*(3), 1046-1053. doi:10.2134/jeq2015.08.0441

Taebi, A., & Droste, R. L. (2004). First flush pollution load of urban stormwater runoff. *Journal of Environmental Engineering and Science, 3*(4), 301-309. doi:10.1139/s04-018

Toran, L., & Grandstaff, D. (2007). Variation of nitrogen concentrations in stormpipe discharge in a residential watershed. *Journal of the American Water Resources Association, 43*(3), 630-641. doi:10.1111/j.1752-1688.2007.00050.x

Wang, Q. G., Li, S. B., Jia, P., Qi, C. J., & Ding, F. (2013). A Review of Surface Water Quality Models. *Scientific World Journal*. doi:Artn 231768

10.1155/2013/231768

Yakub, G. P., Castric, D. A., Stadterman-Knauer, K. L., Tobin, M. J., Blazina, M., Heineman, T. N., . . . Frazier, L. (2002). Evaluation of Colilert and Enterolert defined substrate methodology for wastewater applications. *Water Environment Research, 74*(2), 131-135. doi:Doi 10.2175/106143002x139839

Soil Survey Staff, Natural Resources Conservation Service, United States Department of Agriculture. Soil Survey Geographic (SSURGO) Database 2014. Available online at http://sdmdataaccess.nrcs.usda.gov/. Accessed 03/01/2015.

Tennessee Impaired Waters and TMDL Information. (n.d.). Retrieved September 02, 2016, from https://iaspub.epa.gov/tmdl/attains_state.control?p_state=TN

American Public Health Association, American Water Works Association, and Water Environment Federation (APHA), 1998, 2005. Standard Methods for the Examination of Water and Wastewater, twentyfirst, twentieth ed. American Public Health Association, Alexandria, VA.

Nonpoint source: Urban Areas. (2016, October 17). Retrieved October 27, 2016, from https://www.epa.gov/nps/nonpoint-source-urban-areas

US EPA. (n.d.). Retrieved June 23, 2016, from https://iaspub.epa.gov/waters10/attains_nation_cy.control?p_report_type=T

APPENDIX

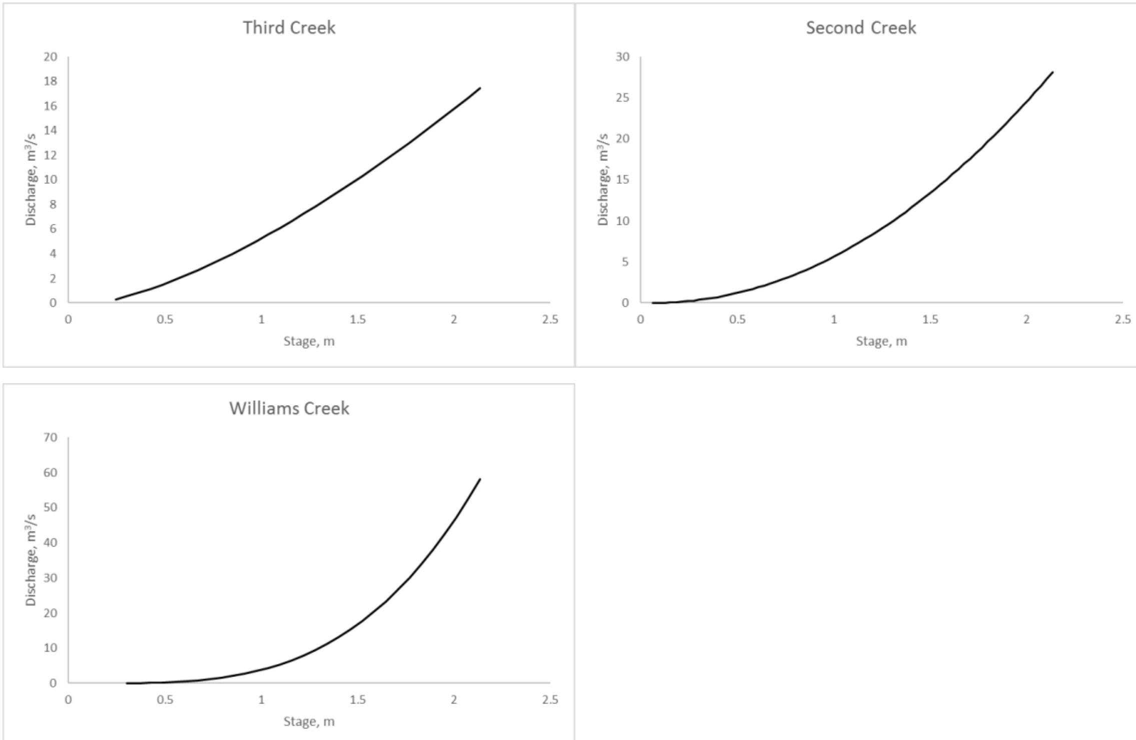


Figure 9. The stage-discharge curves for each in-stream monitored location.

Figure 10. Hydrographs for each Second Creek event that was used in analysis.

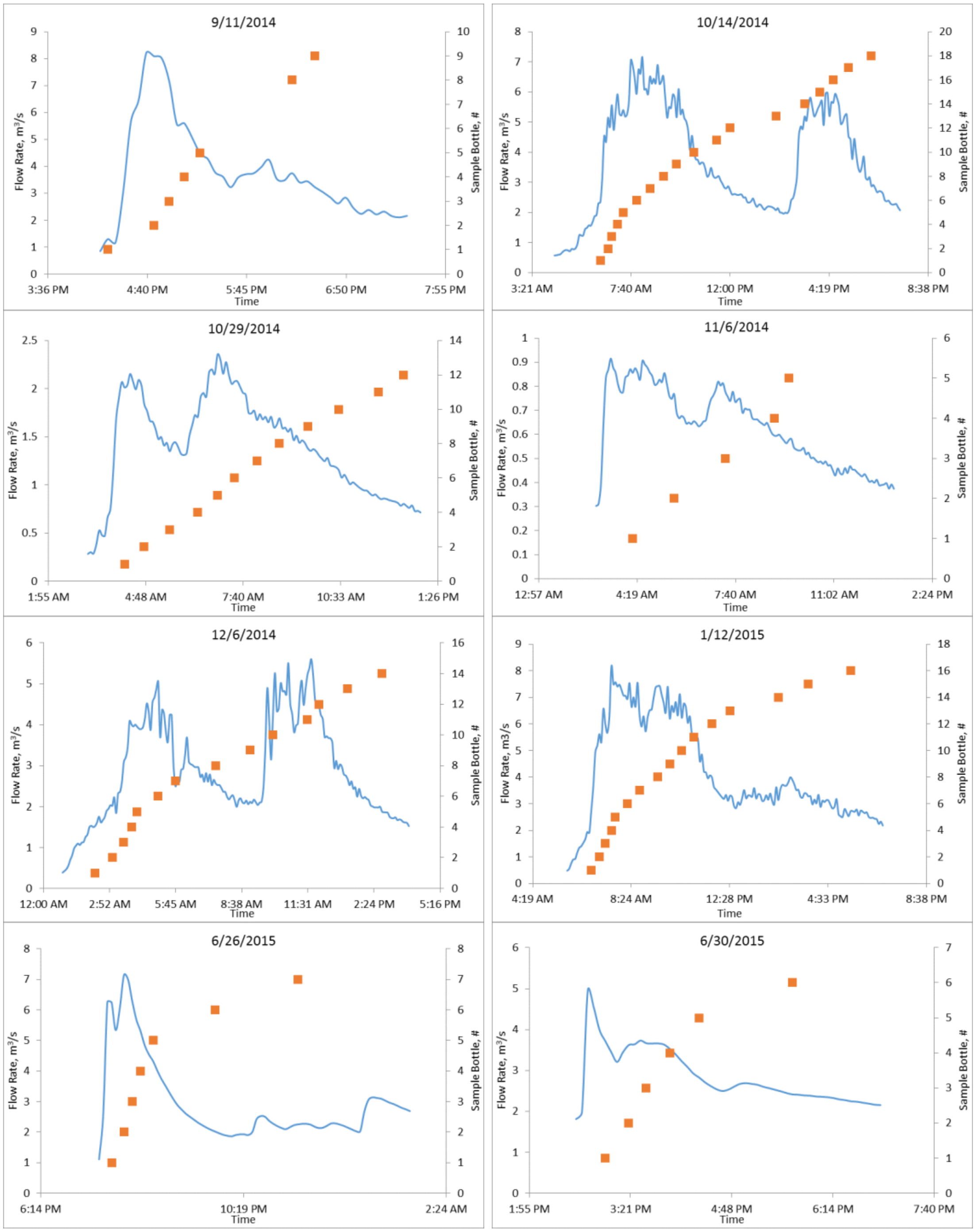


Figure 10 continued

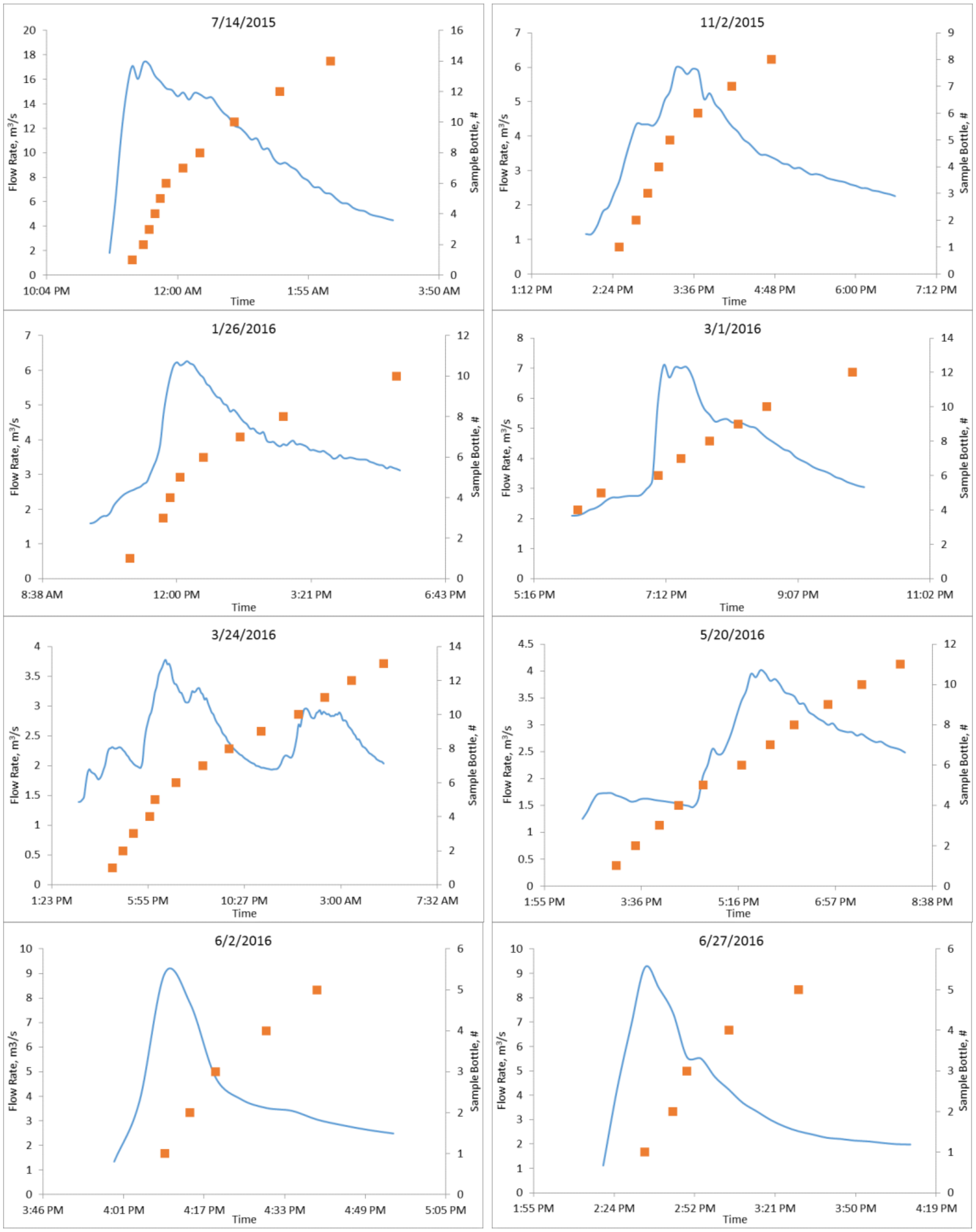


Figure 10 continued

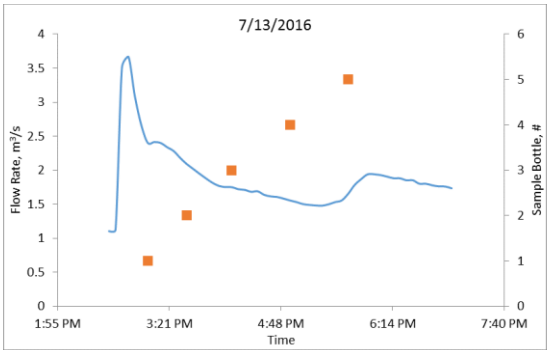


Figure 10 continued

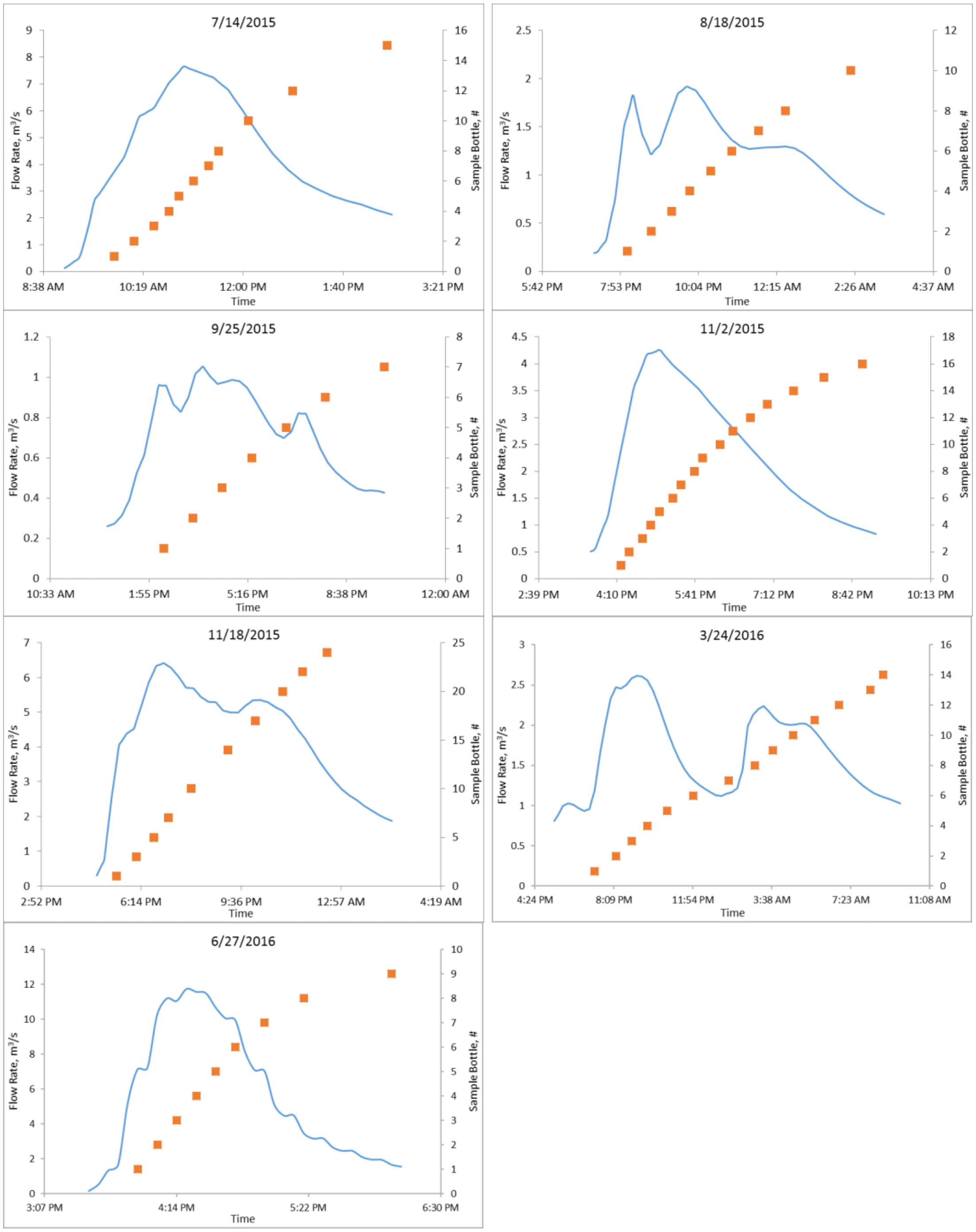


Figure 11. Hydrographs for each Third Creek event that was used in analysis.

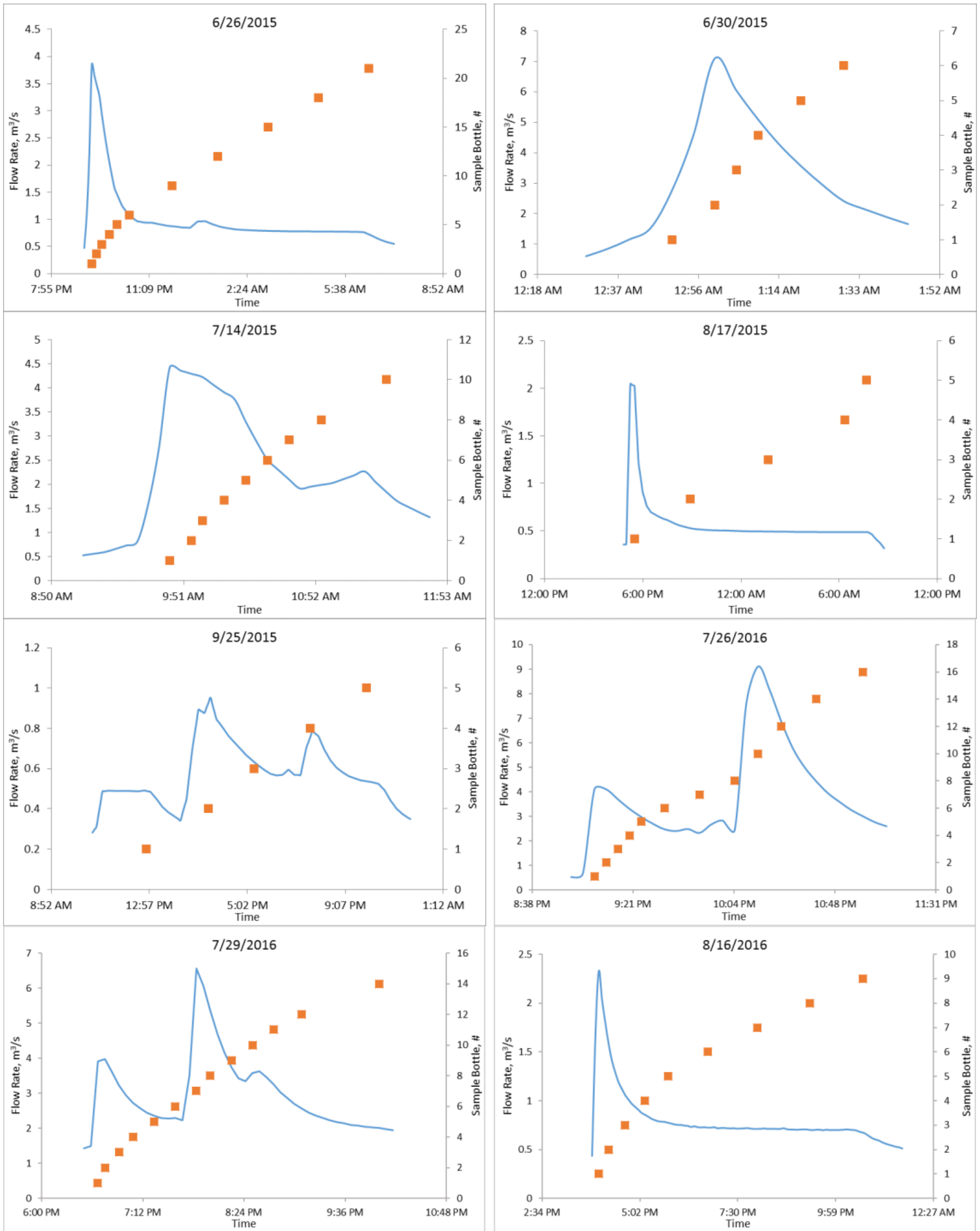


Figure 12. Hydrographs for each Williams Creek event that was used in analysis

VITA

Laurel Elizabeth Christian was born on December 30th, 1992 in Knoxville, Tennessee. She graduated from Central High School in spring of 2011. Upon graduation, Ms. Christian attended the University of Tennessee at Knoxville; where she became employed by Dr. Jon M. Hathaway as an undergraduate research assistant. After graduating Magna Cum Laude from the University of Tennessee with a Bachelor of Science in Civil and Environmental Engineering in May 2015, Ms. Christian had an opportunity to continue her education in graduate school as a graduate research assistant under the guidance of Dr. Hathaway. Ms. Christian will receive her Masters of Science in Environmental Engineering in December of 2016.